LOW-COST ORBIT DETERMINATION USING EXISTING LOW-POWER K-BAND COMMUNICATIONS SIGNALS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Ryan Cean Dougherty

May 2015

LOW-COST ORBIT DETERMINATION USING EXISTING LOW-POWER K-BAND COMMUNICATIONS SIGNALS

Ryan Cean Dougherty, Ph.D.

Cornell University 2015

This work constitutes an effort to examine a new, low-cost orbit determination method for Earth-orbiting satellites which transmit an always-on telemetry signal. Currently, orbit determination for many commercial satellites is performed using two-way ranging *via* 10–15 m high-gain antennas. The cost of orbit determination is driven largely by the capital equipment costs of purchasing the necessary earth stations or leasing time on existing stations. For a typical orbit-raising campaign, the cost to lease earth station time may exceed $100000. The new orbit determination method employs only low-cost, low-gain antennas and general purpose computing equipment. The cost to purchase an entire network of low-cost ground stations, which could be employed to perform orbit determination for many orbit raising campaigns, should be commensurate with the cost of leasing earth station time for a single orbit raising campaign.

The concept of operations of the new, low-cost method includes several ground stations, probably not less than 10 and not more than 20, each equipped with a low-gain antenna for receiving the K_u band spacecraft telemetry signal, down-conversion equipment, digitization equipment, and a general-purpose computer. Each ground station would observe a spacecraft's normal telemetry signal at intervals and derive one or more signal observables to be used in an orbit determination filter or other orbit estimation algorithm. This operational concept differs from the currently-employed concepts in that it does not entail the enormous capital equipment costs that accompany typical two-way ranging techniques. A potential drawback of this method is that since it employs only low-gain antennas, the signal-to-noise ratio is necessarily lower, and the downlink telemetry bit rate is concomitantly lower.

This dissertation provides a signal model for the specific signal used for testing, as well as coarse and fine signal detection algorithms. These algorithms are tested on data gathered from an on-orbit satellite signal. The fine acquisition algorithm provides a method of deriving signal observables usable by an orbit determination algorithm. Finally, an analysis of orbit-determination performance using only the most simple observable that can be derived from the signal detection algorithm is presented.

For a modest number of ground stations, as few as six, real-time orbit-determination accuracy on the order of a kilometer can be expected. Even in the worst cases, simulated position estimation errors climb to only a few kilometers RMS. When more ground stations are employed, the simulated orbit determination accuracy reaches a few tens of meters. This method, or a derivative of it, holds great potential for providing medium-precision orbit determination for a small fraction of the cost of traditional two-way ranging techniques.

BIOGRAPHICAL SKETCH

Ryan Dougherty was born and raised in Sacramento, California. In high school, he enrolled in a program that allowed him to earn his private pilot license concurrently with his high school diploma. Flying was a large part of his life for those four years and swayed his choice of college major from genetic engineering to aerospace engineering. He entered the aerospace engineering program at the University of Southern California in 1999. Sometime during those years in college, his interest shifted from atmospheric flight to spaceflight. He graduated *magna cum laude* in spring of 2003 with a Bachelor of Science degree in aerospace engineering with emphasis on astronautics.

Immediately following graduation, he spent some time in northern Minnesota, near a small town called Ely. He built cabins, roofed sheds, fished in the lake, and trained sled dogs. There were 40 dogs in the dog yard, and he knew them all by name. The winter was cold, but it was an adventure, and he loved it.

After his stint in Minnesota, he came back to civilization and worked for about a year and a half in the Jet Propulsion Laboratory's advanced propulsion group. He did electric propulsion systems engineering for proposed missions, and even one mission that flew a gridded ion engine, DAWN. NASA's direction changed in 2005, and as the most junior member of the group Ryan thought it prudent to move on before he was forced to. Space Systems/Loral hired him into the propulsion engineering group. He worked there for three years before he felt the need to earn an advanced degree. In those three years, Ryan met and married his wife, Alice, who graciously consented to follow him across the country into a purportedly frozen wasteland—if you believed the things you heard in California (though, looking back, Ryan might now describe California as a parched wasteland).

Ryan came to Cornell in 2008, to Professor Mark Psiaki's group. He immediately began working on artificial satellite orbit determination. His special interest was in using existing signals to determine the orbit of satellites already in the sky. This eventually led to a research contract, the results of which you are reading now.

For Elinor, who started,

and Jeeves, who stopped.

ACKNOWLEDGEMENTS

My funding has been largely provided by the wonderful people in the MAE department here at Cornell, through a number of RA-ships and TA-ships. I would like to thank and acknowledge SSL for funding, and former colleagues of mine at SSL, Chris Hoeber, Michel Baylocq, Rick Voreck, and Jeff Stoen, with whom I worked closely during the final phases of my PhD work. I would also like to thank Boeing Defense, Space, and Security for partial support through Contract No. 8CC0789H, which was part of their iGPS project.

The experiment described in the dissertation was carried out on the campus of Echostar Satellite Services in Gilbert, Arizona. We used some of their equipment and some of their people. They deserve a big thanks.

And of course, my advisor, Mark Psiaki, provided invaluable technical insight and advice.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The traditional method of performing orbit determination for commercial communication satellites during orbit raising—the period of time from the moment the satellite is decoupled from the launch vehicle to the time that it reaches its final orbital slot—is to use two-way ranging. A large (10–15 m) parabolic antenna is trained on the satellite, a signal is sent to the satellite, and the satellite sends a reply. Based on either the known delay on the satellite or a carrier phase ambiguity resolution procedure, a two-way range can be calculated. This method of orbit determination is expensive, because it employs a network of high-value Earth stations, and ranging time must be leased on the equipment for the duration of the orbit determination. For a typical orbit raising campaign, which might last 6 to 8 days, the cost may run into the hundreds of thousands of dollars [1]. However, there are certain proposed orbit raising campaigns that would take much longer [2] and cost commensurately more. To enable the business case for these proposed missions, a lower-cost method of providing orbit estimates must be employed.

This dissertation provides the initial proof-of-concept of a low-cost orbit determination method specifically applicable to commercial communication satellites during orbit raising, but more generally applicable to any satellite during any phase of its life. The concept uses a modestly-sized network of passive, low-cost Earth stations. There would be perhaps 10–20 of these stations around the globe, each consisting of on the order of \$5000–\$10000 worth of equipment and requiring power, an internet connection, and a clear view of the sky. The low-cost Earth stations passively detect the always-on spacecraft telemetry signal. Depending on the details of the actual implementation, the telemetry signal would need to be modified to various extents. In the most extreme case, the telemetry signal would broadcast not telemetry bits, but a pseudo-random noise code unique to each satellite. This would be necessary in a situation in which only very-low-gain antennas were observing the

satellite. In a less extreme case, one or more medium-gain antennas, such as standard 30-cm K_u band satellite television antennas, would observe the satellite in addition to any other low-gain antennas. These medium gain antennas would have sufficient gain to read low-rate telemetry data. They would subsequently distribute the known bits to the other ground stations which would post-process recorded RF samples using the known bit stream. In all cases, however, the bit rate would need to be reduced, which could be accomplished readily for most satellite architectures through a simple software change.

From the signal detection, spacecraft carrier Doppler shift would be measured, and each measurement from each Earth station would be sent to a central filter for processing in real time.

The value of this work is in bringing domain knowledge from one area to bear on another. The communication spacecraft community is by nature very conservative and risk-averse. This is entirely understandable, as each satellite costs on the order of a quarter billion dollars and must function without fail for 15 to 20 years. As a result, technology is slow to infuse into flight systems. It must be proven. It must have a track record. And it must have a viable business purpose. For an initial cost of roughly one hundred thousand dollars, this new orbit-determination method could potentially save on the order of one million dollars per spacecraft for extended orbit raising campaigns, and if employed on the higher-volume,lower-cost standard orbit raising campaigns, up to about the cost of the system per spacecraft could be saved in orbit-determination costs. The work embodied in this dissertation is the first step in proving the feasibility of this new approach.

The key element of this proof is an experiment carried out on 11–13 March 2014 in Gilbert, Arizona on the Echostar Satellite Services campus. During the experiment, an actual on-orbit satellite telemetry signal was digitally sampled and recorded on two different receiving paths. On one path, the high-SNR path, a 13-m antenna was used to receive the signal. The resulting carrier power-to-noise power-density ratio (C/N_0) was on the order

of 75 dB-Hz. On the other path, the low-SNR path, a small antenna was used (either a 15-cm standard gain horn, or an omnidirectional antenna, depending on the particular data collection run) to receive the telemetry signal. The C/N_0 for the low-SNR path was about 32 dB-Hz for the SGH and about 11 dB-Hz for the omnidirectional antenna. Acquisition results from the high-SNR path were used as "truth" against which to compare acquisition results from the low-SNR path. There was some question as to whether it was possible to recover the signal from well below the noise floor. The signal is not designed to be received or processed in this way, it is not normally done this way, and its users are not accustomed to such a concept. The experiment proved that it is possible to detect the telemetry signal on small, inexpensive ground stations.

The experiment provided several successful signal detections. Based on the measurement noise variance anticipated from the signal detections, orbit determination analyses show that sub-kilometer orbit knowledge should be attainable with as few as ten ground stations, and sub-100-meter orbit knowledge should be reachable with 20.

The ensuing chapters present the signal-modeling work, the acquisition algorithms, results of acquisitions on real data, and results of orbit determination simulations using simulated observables such as would be obtainable using low-gain antennas. The results contained in this work establish that it is feasible to detect the telemetry signal using low-cost ground stations, that it is possible to measure at least carrier Doppler shift to sufficient precision to be useful in orbit determination, and that orbits derived from these measurements would be of high enough precision to be useful.

The following is an overview of the remainder of this dissertation.

Chapter 2 provides an in-depth development of a mathematical model of the telemetry signal. The signal model is the foundation on which the entire low-cost orbit determination method is built. All later chapters use this model.

Chapter 3 describes the coarse acquisition method and its application to real data from an on-orbit satellite. The results of this chapter are the key to the proof of the low-cost orbit determination concept. This chapter demonstrates signal acquisition for real satellite signals recorded using representative equipment. The most significant technical risk at the outset of this work was whether the signal would be detectable at all. This chapter presents proof that the signal is detectable, even at very low carrier power to noise power density ratios (hereinafter referred to as simply "carrier-to-noise ratio" or abbreviated C/N_0).

Chapter 4 describes the fine acquisition algorithm and the signal observables. The fine acquisition algorithm takes the results from the coarse acquisition and optimizes the signal parameters to recover the most signal power. A portion of the signal parameter vector describes the carrier Doppler shift over the detection interval. The Doppler shift at the center of the fine acquisition interval is used as a signal observable to be handed to the orbit determination filter. The chapter concludes with an analysis of the observable and its quality for the different acquisition scenarios described in Chapter 3.

Chapter 5 presents a simplified orbit determination analysis for the proposed system. The orbit-determination analysis uses simulated orbits and measurements with realistic levels of process noise and measurement noise. The orbit and measurement models are simple but adequate for the purpose at hand—that is, to demonstrate rough order-of-magnitude performance and to identify trends. Several Earth-station locations and different combinations of locations are studied. Two levels of measurement noise are studied, each corresponding roughly to a level of noise expected from the low-gain antenna systems that have been used to generate the results of the acquisition experiments presented in Chapter 3. The results of the analysis are presented in a table and several figures.

Finally, Chapter 6 presents a summary of the dissertation, including a discussion of the main contributions of the dissertation. This chapter also presents some ideas regarding possible ways to proceed in order to apply the work in this dissertation to an operational

orbit determination system.

CHAPTER 2

SIGNAL DESCRIPTION

The goal of the work in this dissertation is to provide foundational elements of the low-cost orbit determination method described in the introduction. The orbit determination method uses passive measurements of signal observables of the always-on satellite downlink telemetry signal used on communication satellites, though any always-on signal could be employed. Signal detection and the consequent signal measurements are the most difficult aspect of the low-cost orbit determination concept.

A signal model is the foundation of the signal detection algorithm. This chapter sets forth a signal model applicable to the particular class of satellite telemetry signal investigated for the low-cost orbit determination concept.The general structure of the detection algorithm could be adapted for most continuously broadcast signals, though a different signal model would need to be developed.

2.1 Signal Model

The class of telemetry signal under investigation is broadcast at a minimum power level of roughly 10 dB-W effective isotropic radiated power (EIRP) in the K_u band, but the broadcast power level varies by satellite. The satellite used as a test case broadcast at roughly 20 dB-W EIRP. By the time the signal arrives at the ground, the received carrier-to-noise ratio for an omnidirectional ground station antenna has fallen to a worst-case estimated 0.17 dB-Hz. A link budget for a worst-case satellite and two budgets for the test-case satellite with different receiving antennas are presented in Table 2.1. The link budgets presented in the table contain a number of lines, each of which provides the assumed gain, loss, or parameter value for each of the three columns. For instance, the first line of the table shows that the

EIRP for the worst-case scenario is 10 dB-W and for the other two scenarios is 20 dB-W. The differences between the columns which result in different C/N_0 are mainly the EIRP and the receiving antenna gain—one of the test-case columns assumes an omnidirectional antenna and the other assumes a standard gain horn (SGH). The worst-case range is also considered in the first column. The last line of the table shows the C/N_0 for each scenario. It is this parameter that describes how far buried in the noise the signals are, and thus how difficult they are to detect, acquire, and track. For comparison, GPS signals, which are also well below the noise floor, are considered weak at C/N_0 of about 35 dB-Hz and are typically acquired and tracked at C/N_0 of 40–50 dB-Hz[3]. Extracting useful information from the telemetry signal buried well below the noise floor is a challenging problem. It is not impossible, however, as will be demonstrated in the coming chapters.

Table 2.1 – Link budgets for the test satellite and for a "worst-case" satellite.

Quantity	Units	Worst Case	Test Sat., Omni	Test Sat., SGH
Transmitter EIRP	dB-W	10	20	20
Range	km	41679	38069.83	38069.83
Path Loss	dB	163.39	162.604	162.604
Multipath Loss	dB	3	3	3
Atmospheric Loss, Rain	dB	3.1	3.1	3.1
Atmospheric Loss, Clear	dB	1.5	1.5	1.5
Power Density at Rcvr	dB-W/m^2	-160.990	-150.204	-150.204
Frequency	GHz	12.6985	12.6985	12.6985
Wavelength	m	0.024	0.024	0.024
Receiving Antenna Gain	dB	3	3	24
Rcx Antenna Eff. Area	dB-m^2	-40.531	-40.531	-19.531
Receiver Internal Losses	dB	1	1	1
Received Signal Power	dB-W	-202.521	-191.734	-170.734
Ext. Eff. Noise Temp.	K	90	90	90
LNA Eff. Noise Temp.	K	300	300	300
C/N_0	dB-Hz	0.169	10.956	31.956

The telemetry signal [4] consists of Manchester-encoded data that has been binary phase shift keyed (BPSK) onto a 48 kHz sinusoidal subcarrier signal. The data is arranged into telemetry frames of 2048 bits at a rate of 4800 bits per second—ten subcarrier cycles elapse per data bit. The bits are Manchester encoded, *i.e.* the bit stream is modulated by a

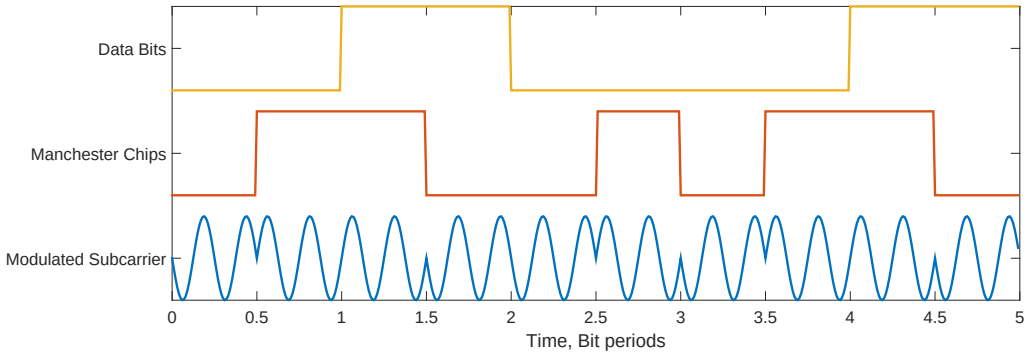


Figure 2.1 – An example string of data bits is represented by the upper trace (yellow). The bits shown are [-1, 1, -1, -1, 1]. The second trace (red) shows the relationship of the Manchester chips to the data bit values. Each bit is represented by two opposite-polarity chips whose order ([1, -1] or [-1, 1]) is determined by the value of the bit. The bottom trace (blue) shows the sinusoidal subcarrier wave modulated with the Manchester chips. Note that the subcarrier shown has a frequency of four times the bit rate, whereas the actual carrier has a frquency of ten times the bit rate. This lower-rate subcarrier more clearly illustrates the relationships than the full-rate subcarrier.

square wave with period equal to the bit period so that each bit is represented by two chips. There is a guaranteed chip transition at the center of each bit, because the square wave changes value at the midpoint of each bit. For instance, a '1' bit is represented with the two chips '1' and '-1' while a '-1' bit is represented with the two chips '-1' and '1'. The BPSK modulation causes a 180° subcarrier phase shift to occur each time the product of the Manchester chip and the data bit changes sign. The relationship of the data bits, the Manchester chips, and the subcarrier is shown in Figure 2.1. The subcarrier phase modulates the RF carrier signal. The phase modulation causes the carrier phase to be modified by a number of radians proportional to the subcarrier's instantaneous amplitude. The carrier and subcarrier are generated by different clocks, referred to as spacecraft clock A and spacecraft clock B, respectively. The chip transitions are phase coherent with the subcarrier, occurring at zero-crossings of the subcarrier. This description motivates a simple model of the signal at the output of the RF front end:

$$y = A \exp\left(j\left[\omega_A t_A + \phi_A + A_{PM} B(t_B - t_{B0}) \sin\left(\omega_B (t_B - t_{B0})\right) - \omega_{mix} t\right]\right) + n \qquad (2.1)$$

where A is the amplitude of the signal at the output of the RF front end (*i.e.* after mixing to the intermediate frequency), j is the imaginary unit value $\sqrt{-1}$, ω_A is the nominal carrier frequency in rad/s, ω_B is the nominal subcarrier frequency in rad/s, t_A is the time of transmission in seconds according to the erroneous spacecraft clock A, t_B is the time of transmission in s according to the erroneous spacecraft clock B, ϕ_A is the arbitrary initial phase of the signal, t_{B0} is a time offset that serves to time-align the subcarrier and its code payload with the RF carrier, A_{PM} is the amplitude of the phase modulation of the RF carrier, $B(t_B - t_{B0})$ is the value of the Manchester chip multiplied by the data bit at time $t_B - t_{B0}$ (both the Manchester chips and the data bits are represented as values of either +1 or -1, so this quantity takes on values of $\pm$ 1), ω_{mix} is the nominal mixing frequency off the RF front end in rad/s, t is the time in seconds according to the erroneous mixing (or receiver) clock, and $n = n_I + jn_Q$ is zero-mean, white, Gaussian complex-valued measurement noise.

In the preceding description, the RF carrier is modulated by only one subcarrier signal. However, the telemetry signal can actually be modulated by as many as three completely independent subcarriers with their own, independent data streams. This modulation is sometimes referred to as "multi-tone angle modulation" [5]. The concept of operations of the low-cost orbit determination scheme would preclude the use of either of the other two additional subcarrier signals due to the added complexity of acquiring the additional signal. During the proof-of-concept experiment, described in detail in Chapter 3, one additional subcarrier signal was being broadcast by the satellite. In order to model multiple subcarriers it is necessary to modify Eq. (2.1)

$$y = A\exp\left(j\left[\omega_A t_A + \phi_A + A_{PM,S}\sum_{s=1}^{S}\left\{B_s(t_B - t_{Bs0})\sin\left(\omega_{Bs}(t_B - t_{Bs0})\right)\right\} - \omega_{mix}t\right]\right) + n \tag{2.2}$$

The subscript s identifies the subcarrier by number. The positive integer S is the total number of subcarriers in operation. Each of the subcarriers has its own subcarrier frequency ω_{Bs},

its own bit stream B_s, and its own code start time t_{Bs0}. It is possible that for some systems, all t_{Bs0} might be equal because the subcarriers are all generated by the same clock and the data bits are synchronized. However, in the case of the particular satellite investigated, the start times of the bits were only guaranteed to start at a particular subcarrier phase, with no guarantee about the bit phase between subcarriers. The phase modulation amplitude depends on the number of subcarriers operating by $A_{PM,S} = 1/\sqrt{S}$. In the remainder of the discussion of the signal model, only one subcarrier will be discussed.

The signal model in Eq. (2.1) is an intuitive description of the signal at the output of the front end, but it must be reworked before being useful for acquiring the signal in digital samples. A first step is to rewrite the model in terms of the receiver clock time. The following relationships among the various clock times are necessary.

$$t_A = T_{AB} + \delta t_A \tag{2.3}$$

$$T_{AB} = T - \frac{\rho}{c} \tag{2.4}$$

$$t_B = T_{AB} + \delta t_B \tag{2.5}$$

$$T = t - \delta t_R \tag{2.6}$$

These equations relate the erroneous spacecraft clock times, denoted by t_A and t_B, to the erroneous receiver clock time, denoted by t, through the true time of transmission denoted T_{AB} and reception denoted T and the respective time-dependent clock errors for each of the three clocks, δt_A, δt_B, δt_R. The variable ρ represents the true range between the spacecraft and receiver at the time of signal transmission, and c is the speed of light in vacuum. These equations can be combined to form two useful relationships.

$$t_A = t - \delta t_R - \frac{\rho}{c} + \delta t_A \quad \text{and} \tag{2.7}$$

$$t_B = t - \delta t_R - \frac{\rho}{c} + \delta t_B \tag{2.8}$$

That is, the erroneous spacecraft clock times at the instant of transmission can be related to the erroneous receiver clock time at the instant of reception. It is also convenient to

define the intermediate frequency as $\omega_{IF} = \omega_A - \omega_{mix}$ so that terms in the exponential that multiply the receiver clock time can be collected.

$$y = A\exp\left(j\left[\omega_{IF}t + \left(\omega_A(-\delta t_R - \frac{\rho}{c} + \delta t_A) + \phi_A\right)\right.\right. \\ \left.\left. + A_{PM}B(t - \delta t_R - \frac{\rho}{c} + \delta t_B - t_{B0})\sin\left(\omega_B(t - \delta t_R - \frac{\rho}{c} + \delta t_B - t_{B0})\right)\right]\right) + n \tag{2.9}$$

The model is now written in terms of quantities that can be measured or modeled, but this is still not convenient for signal detection. One can use a polynomial to approximate the expression in the square brackets. A quadratic might suffice in certain situations, *i.e.*,

$$\omega_A(-\delta t_R - \frac{\rho}{c} + \delta t_A) + \phi_A \approx \frac{\alpha_{Ak}}{2}(t - t_k)^2 + \omega_{DAk}(t - t_k) + \phi_{Ak} \tag{2.10}$$

The epoch time, t_k, defines the beginning of the period of applicability of the polynomial fit, and α_{Ak}, ω_{DAk}, and ϕ_{Ak} are the quadratic, linear, and constant polynomial coefficients, respectively. The subscript k indicates that these quantities apply at the epoch time t_k, which would typically be the time of the first sample used in a signal detection. Another polynomial can be used to approximate the terms related to spacecraft clock B error. For example, a quadratic model would take the form:

$$\omega_B(-\delta t_R - \frac{\rho}{c} + \delta t_B - t_{B0}) \approx \frac{\alpha_{B0}}{2}(t - \tau_{B0})^2 + \omega_{DB0}(t - \tau_{B0}) - \omega_B\tau_{B0} \tag{2.11}$$

In this polynomial approximation, τ_{B0} is the erroneous receiver clock time at which the phase of the subcarrier goes to zero, which is the same as the reception time of the subcarrier and bit stream features that were transmitted at erroneous transmitter clock time t_{B0}. Specifying τ_{B0} amounts to specifying the unknown initial subcarrier phase. The other new quantities, α_{B0} and ω_{DB0}, are the quadratic and linear polynomial coefficients. It is conceivable that some advantage could be gained by modeling separately the ρ/c term which is common to both of these polynomial approximations, Eqs. (2.10) and (2.11). However, experience with

signal detection shows that it is easier to treat the carrier and the subcarrier independently. This allows the carrier and subcarrier to be acquired independently. Now, the i^{th} signal sample can be written

$$\begin{aligned} y_i = A\exp\Bigg(j\Bigg[\omega_{IF}t_i + \left[\frac{\alpha_{Ak}}{2}(t_i-t_k)^2 + \omega_{DAk}(t_i-t_k) + \phi_{Ak}\right] \\ + A_{PM}B\left(\frac{\alpha_{B0}}{2\omega_B}(t_i-\tau_{B0})^2 + \left(1+\frac{\omega_{DB0}}{\omega_B}\right)(t_i-\tau_{B0})\right) \\ \sin\left(\frac{\alpha_{B0}}{2}(t_i-\tau_{B0})^2 + (\omega_B+\omega_{DB0})(t_i-\tau_{B0})\right)\Bigg]\Bigg) + n_i \end{aligned} \tag{2.12}$$

The seven unknowns in this form of the model are A, α_{Ak}, ω_{DAk}, ϕ_{Ak}, α_{B0}, ω_{DB0}, and τ_{B0}. Using these polynomial approximations will simplify the process of signal acquisition.

At this point it is also convenient to define some shorter names for the quantities of interest. Define

$$\alpha_i = \omega_{IF}t_i + \frac{\alpha_{Ak}}{2}(t_i-t_k)^2 + \omega_{DAk}(t_i-t_k) \tag{2.13}$$

and

$$\begin{aligned} \beta_i = A_{PM}B\left(\frac{\alpha_{B0}}{2\omega_B}(t_i-\tau_{B0})^2 + \left(1+\frac{\omega_{DB0}}{\omega_B}\right)(t_i-\tau_{B0})\right)\sin\Big(\frac{\alpha_{B0}}{2}(t_i-\tau_{B0})^2 \\ + (\omega_B+\omega_{DB0})(t_i-\tau_{B0})\Big) \end{aligned} \tag{2.14}$$

This significantly shortens the representation of the model

$$y_i = A\exp\left(j\left(\phi_{Ak}+\alpha_i+\beta_i\right)\right) + n_i \tag{2.15}$$

In Eq. (2.15), six of the unknown quantities enter the model non-linearly. However, it is possible to re-write the model such that only five of the unknowns enter non-linearly. The following definitions achieve this.

$$X = A\cos(\phi_{Ak}) \tag{2.16a}$$

$$Y = -A\sin(\phi_{Ak}) \tag{2.16b}$$

Next, these definitions can be substituted into Eq. (2.15) to yield:

$$y_i = (X - jY) \exp\left(j\left(\alpha_i + \beta_i\right)\right) + n_i \tag{2.17}$$

The unknowns in this form of the signal model are X, Y, α_{Ak}, ω_{DAk}, α_{B0}, ω_{DB0}, and τ_{B0}. There are seven unknowns, but two of them, X and Y, enter the model linearly. These signal model parameters will be used in the remainder of this dissertation.

Before moving on, it is worthwhile to review the meanings of the seven unknowns. The initial carrier Doppler shift is modeled by ω_{DAk}, and its rate of change by α_{Ak}. These quantities are expressed in units of rad/s and rad/s^2. The carrier Doppler shift is due to satellite and receiver clock errors and relative motion of satellite and receiver. The satellite clock errors should dominate this term for satellites in geostationary orbits—in the worst case scenario, the two terms will be of roughly equal magnitude—but in the low-altitude portion of GTO orbits, the effects of relative motion will dominate.[1] The initial subcarrier Doppler is modeled by ω_{DB0}, and its rate of change by α_{B0}. The receiver clock time at which the code starts is modeled by τ_{B0}. The two remaining parameters, X and Y, model the signal amplitude and initial carrier phase via the relations in Eqs. (2.16a) and (2.16b). Note that the signal parameters X and Y may be re-optimized for each coherent accumulation interval[2] m, of M intervals, resulting in $2M$ parameters X_m and Y_m. As a result, there may be $2M + 5$ signal parameters. A $2M + 5$ dimensional parameter space may seem daunting, but the search space can be effectively cut down to five dimensions, as will be shown in the next section.

[1] Relative motion may produce frequency shifts of up to roughly 600 kHz in the low-altitude portion of GTO orbits and up to 30 kHz for a geostationary satellite which swings across its entire half-degree orbital slot. A clock accurate to 1 part-per-million would produce frequency errors of 18 kHz at the top of the K_u band.

[2] Coherent accumulation is described in Sec. 3.1

2.2 Estimation of Linear Parameters

The subject of this section is an aspect of signal processing, but is included in this chapter on the signal model. The reason for this is that the particular form of the signal model necessitates special considerations during signal processing. Specifically, the signal model allows for phase and amplitude discontinuities at the boundaries of coherent accumulation intervals. These discontinuities are accounted for by the $2M$ independent parameters X_m and Y_m for $m = 1 \ldots M$.

These X_m and Y_m parameters enter the signal model linearly and may be estimated in a linear estimation step prior to the much more difficult nonlinear estimation of the other five signal parameters. This effectively reduces the signal model parameter search space to a five-dimensional space from $2M + 5$ dimensions. As such, it is appropriate to discuss this aspect of estimation in connection with the description of signal model.

The discussion of this linear estimation process can begin by considering a single coherent accumulation interval. The two linear parameters for this interval may be optimally estimated, in the minimum mean squared error sense, separately from the remaining five model parameters. The process begins by referring back to Eq. (2.17), and recognizing that the signal sample can be represented as the sum of a real and an imaginary part.

$$y_i = (X_m - jY_m)\left[\cos(\alpha_i + \beta_i) + j\sin(\alpha_i + \beta_i)\right] + n_{Ii} + jn_{Qi} \tag{2.18}$$

In the preceding equation, the complex-valued noise term n_i has been replaced by its real and imaginary components, n_{Ii} and n_{Qi}. By defining $a_i = \cos(\alpha_i + \beta_i)$ and $b_i = \sin(\alpha_i + \beta_i)$ and after some minor algebra,

$$\begin{aligned} y_i &= (a_i X_m + b_i Y_m + n_{Ii}) + j\,(b_i X_m - a_i Y_m + n_{Qi}) \\ &= (y_{Ii} + n_{Ii}) + j(y_{Qi} + n_{Qi}) \end{aligned} \tag{2.19}$$

The second line of this equation uses the definition of the in-phase and quadrature components of the signal sample as $y_{Ii} = a_i X_m + b_i Y_m$ and $y_{Qi} = b_i X_m - a_i Y_m$. Consider some unspecified coherent accumulation interval over which carrier phase is assumed to be continuous with samples numbered 1 through N inside that interval. The in-phase and quadrature measurement vectors are

$$\boldsymbol{y}_I = \left[y_{I1},\, y_{I2},\, \ldots,\, y_{Ii},\, \ldots,\, y_{IN} \right]^{\mathrm{T}} \tag{2.20}$$

$$\boldsymbol{y}_Q = \left[y_{Q1},\, y_{Q2},\, \ldots,\, y_{Qi},\, \ldots,\, y_{QN} \right]^{\mathrm{T}} \tag{2.21}$$

Each element of $\boldsymbol{y}_I$ or $\boldsymbol{y}_Q$ corresponds to a single measurement sample received at time t_i in the interval under consideration. In the general case, of the m^{th} out of M coherent integration intervals, $t_m < t_i \leq t_{m+1}$ for each i in $N(m-1)+1 < i \leq N(m-1)+N$ and where t_m is the time of the last sample in coherent accumulation interval $m-1$ and t_{m+1} is the time of the last sample in coherent accumulation interval m. That is, the m and i subscripts count different quantities—m subscripts refer to the coherent accumulation interval number and i subscripts refer to the raw sample count. The importance of coherent accumulation intervals will be described in Chapter 3. If all a_i, b_i, n_{Ii}, and n_{Qi} are stacked into vectors as were the in-phase and quadrature measurements, all copies of Eq. (2.19) for all $t_m < t_i \leq t_{m+1}$ can now be written in two vector equations as

$$\boldsymbol{y}_I = \boldsymbol{a} X_m + \boldsymbol{b} Y_m + \boldsymbol{n}_I \tag{2.22a}$$

$$\boldsymbol{y}_Q = \boldsymbol{b} X_m - \boldsymbol{a} Y_m + \boldsymbol{n}_Q \tag{2.22b}$$

In matrix form, these two equations are

$$\begin{bmatrix} \boldsymbol{y}_I \\ \boldsymbol{y}_Q \end{bmatrix} = \begin{bmatrix} \boldsymbol{a} & \boldsymbol{b} \\ \boldsymbol{b} & -\boldsymbol{a} \end{bmatrix} \begin{bmatrix} X_m \\ Y_m \end{bmatrix} + \begin{bmatrix} \boldsymbol{n}_I \\ \boldsymbol{n}_Q \end{bmatrix} \tag{2.23}$$

The optimal least-squares X_m and Y_m can be found from standard least-squares theory. That is, the unweighted sum of the squared errors in Eq. (2.23) will be minimized, *i.e.*,

$$J(X_m, Y_m) = \frac{1}{2} \left(\boldsymbol{n}_I^{\mathrm{T}} \boldsymbol{n}_I + \boldsymbol{n}_Q^{\mathrm{T}} \boldsymbol{n}_Q \right), \tag{2.24}$$

after determining $\boldsymbol{n}_I$ and $\boldsymbol{n}_Q$ as functions of X_m and Y_m by solving for them using Eq. (2.23). Start by multiplying both sides of Eq. (2.23) by the transpose of the two-by-two block matrix on the right-hand side of Eq. (2.23).

$$\begin{bmatrix} \boldsymbol{a}^{\mathrm{T}}\boldsymbol{y}_I + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}_Q \\ \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}_I - \boldsymbol{a}^{\mathrm{T}}\boldsymbol{y}_Q \end{bmatrix} = \begin{bmatrix} (\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b}) & (\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} - \boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}) \\ (\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a} - \boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}) & (\boldsymbol{b}^{\mathrm{T}}\boldsymbol{b} + \boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}) \end{bmatrix} \begin{bmatrix} X_m \\ Y_m \end{bmatrix} + \begin{bmatrix} \boldsymbol{a}^{\mathrm{T}}\boldsymbol{n}_I + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{n}_Q \\ \boldsymbol{b}^{\mathrm{T}}\boldsymbol{n}_I - \boldsymbol{a}^{\mathrm{T}}\boldsymbol{n}_Q \end{bmatrix} \tag{2.25}$$

Following the vector inner products, each entry of Eq. (2.25) is scalar. Recalling that a_i and b_i represent cosine and sine, the diagonal entries of the 2×2 matrix on the right-hand side of the equation are each N and the off-diagonal entries are zero. Taking the expected value of the equation affects only the last term, as it is the only random term. Since the expected value of the sum is the sum of the expected values, and since $\mathrm{E}[n_{Ii}] = 0$ and $\mathrm{E}[n_{Qi}] = 0$, the result of taking the expected value of Eq. (2.25) is

$$\begin{bmatrix} \boldsymbol{a}^{\mathrm{T}}\boldsymbol{y}_I + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}_Q \\ \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}_I - \boldsymbol{a}^{\mathrm{T}}\boldsymbol{y}_Q \end{bmatrix} = \begin{bmatrix} N & 0 \\ 0 & N \end{bmatrix} \begin{bmatrix} X_m \\ Y_m \end{bmatrix} \tag{2.26}$$

Or, after inverting the 2×2 matrix on the right-hand side of Eq. (2.26),

$$\hat{X}_m = \frac{1}{N}\left(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{y}_I + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}_Q\right) \tag{2.27a}$$

$$\hat{Y}_m = \frac{1}{N}\left(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}_I - \boldsymbol{a}^{\mathrm{T}}\boldsymbol{y}_Q\right) \tag{2.27b}$$

The optimal values of X_m and Y_m, $\hat{X}_m$ and $\hat{Y}_m$, minimize $J(X_m, Y_m)$ in Eq. (2.24). That is, the final minimized cost is

$$J(\hat{X}_m, \hat{Y}_m) = \frac{1}{2}\left(\boldsymbol{y}_I^{\mathrm{T}}\boldsymbol{y}_I + \boldsymbol{y}_Q^{\mathrm{T}}\boldsymbol{y}_Q\right) - \frac{1}{2N}\left((\boldsymbol{a}^{\mathrm{T}}\boldsymbol{y}_I + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}_Q)^2 + (\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}_I - \boldsymbol{a}^{\mathrm{T}}\boldsymbol{y}_Q)^2\right) \tag{2.28}$$

Note that $\hat{X}_m$ and $\hat{Y}_m$ constitute accumulations for standard in-phase and quadrature baseband mixing, also known as matched filtering. The optimal values of the two parameters X_m and Y_m are determined by the values of the samples and the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, which in turn depend on the five remaining signal parameters.

These parameters $\hat{X}_m$ and $\hat{Y}_m$ need not be considered explicitly in a signal detection grid search. However, they must be computed in order to calculate the residuals necessary for a nonlinear optimization routine to optimize the other signal parameters. X_m and Y_m must be optimized for each coherent accumulation interval. This amounts to allowing the carrier amplitude and phase to have discontinuities at coherent accumulation interval boundaries. The carrier Doppler shift will remain constant (or ramp according to the carrier Doppler rate), but phase is allowed to have jumps. This is non-physical, but allows the modeled carrier phase over each interval to match the true carrier phase more closely than if model carrier phase were forced to be continuous over all intervals. This situation is illustrated in Figure 2.2. In the figure, the solid blue line represents a true carrier phase error history, the dash-dotted red line is a quadratic fit, and the dashed green curve with discontinuities is a quadratic fit with five separate phase biases which apply over the five separate segments. The fit of the curve with discontinuities is better than that without.

Following the linear estimation step described in this section, there are five signal model parameters which must be optimized in a non-linear estimation procedure. The next section describes the global grid search which results in a rough guess at the optimal values of the five remaining signal parameters.

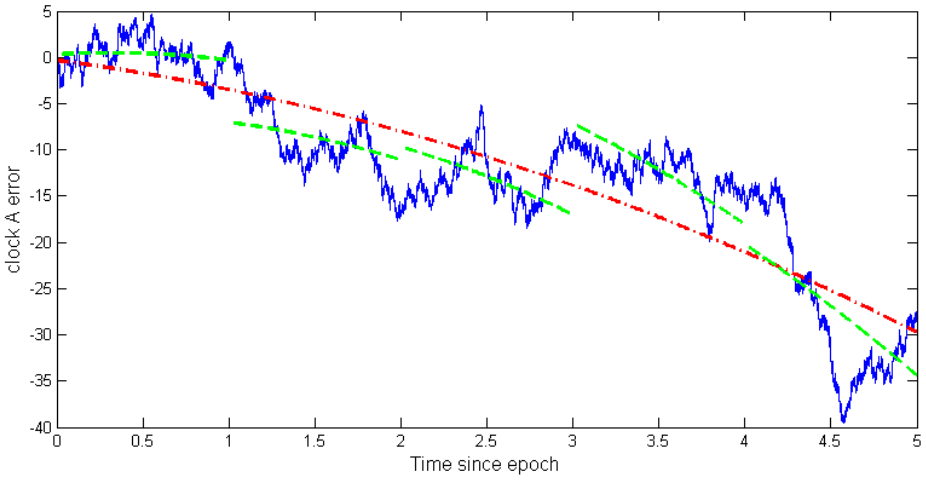


Figure 2.2 – An example Clock A phase error history (solid blue line) with arbitrary units is not well-modeled by a quadratic (red dash-dotted line). In order to fit the random walk of the phase error better, the quadratic is broken into many pieces with their own vertical offsets (green dashed line segments).

CHAPTER 3

COARSE ACQUISITION

The most challenging aspect of the low-cost orbit determination concept is signal detection and acquisition. Once the signal is detected and its observables measured using acquisition-type calculations, it is relatively straightforward to feed the observables to a filter or batch algorithm tailored for orbit determination. The process of acquiring the satellite telemetry signal has been divided into two phases, coarse and fine acquisition. This chapter delineates the coarse acquisition method that has been developed and the results of applying the method to data gathered in the field from an on-orbit satellite. During coarse acquisition, a rough estimate of the signal parameters is found. Fine acquisition is covered in Chapter 4. During fine acquisition, the signal parameters from the coarse acquisition are optimized.

The signal detector employed in this work is a coherent detector with post-detection non-coherent accumulation [3] [6], as opposed to purely coherent or purely non-coherent techniques [7]. Coherent detection relies on being able to construct a replica signal. If the signal contains unknown data bits or error correction symbols, constructing a replica complete with these unknown bits may be prohibitively difficult. This means that no (or very few) unknown data bits can be transmitted during the duration of the signal detection. However, The advantage of coherent detection is that it allows detection of the signal at very low C/N_0, below that at which it is possible to accurately decode the data bits. These considerations may be important for systems engineering trades of an orbit determination system based on this signal detection scheme.

The coarse acquisition phase of the signal detection consists of running the detector at each point in a grid over signal parameter space. At each point in parameter space, one or more coherent accumulations are computed followed by a non-coherent accumulation. The process of computing these accumulations is described in the next sections of this chap-

ter. Simply put, an accumulation can be through of as the summation of the samplewise multiplication of the sampled incoming signal with a receiver-generated replica signal. The number and duration of these accumulations is determined by the system engineer based on C/N_0 and the desired Doppler accuracy. The resulting accumulations are normalized and compared to a detection threshold. The normalized accumulation is referred to as the detection test statistic. A detection test statistic above the threshold is statistically unlikely to result from accumulating only noise. In this case, a signal is declared to have been detected at the point in parameter space that produced the accumulation. It is likely that any reasonable search grid will contain a region, rather than a single point, that produces positive signal detections due to correlations between neighboring points. If the search grid has been chosen appropriately, the set of parameters that produced the highest accumulation in the region will be nearest the global maximum, and it is that maximum that the fine acquisition procedure will find. The remainder of this chapter provides the details of the coarse acquisition algorithm and discusses the results of applying it to data gathered from a spacecraft currently on-orbit.

3.1 Calculating Accumulations

The coherent accumulations are computed by mixing the signal samples with samples of a signal replica and summing the results into in-phase and quadrature coherent accumulations. The m^{th} in-phase and quadrature coherent accumulations are

$$I_m(\hat{S}_k) = \sum_{i=(m-1)N+1}^{mN} \hat{a}(\hat{S}_k, i) y_{Ii} + \hat{b}(\hat{S}_k, i) y_{Qi} \tag{3.1}$$

$$Q_m(\hat{S}_k) = \sum_{i=(m-1)N+1}^{mN} \hat{b}(\hat{S}_k, i) y_{Ii} - \hat{a}(\hat{S}_k, i) y_{Qi} \tag{3.2}$$

In the two preceding equations, $I_m(\hat{S}_k)$ and $Q_m(\hat{S}_k)$ are the m^{th} in-phase and quadrature coherent accumulations given signal model parameter set $\hat{S}_k = \{\hat{\alpha}_{Ak}, \hat{\omega}_{DAk}, \hat{\alpha}_{B0}, \hat{\omega}_{DB0}, \hat{\tau}_{B0}\}$ which applies over the entire non-coherent accumulation interval. Each coherent accumulation involves a summation over N samples, indexed with i, which have been mixed with replica signal samples $\hat{a}(\hat{S}_k, i)$ and $\hat{b}(\hat{S}_k, i)$. These are the same a_i and b_i that appeared earlier, but the circumflex ($\hat{}$) has been added to denote that these are estimates, and the dependence of the estimates on the signal model parameter estimates has been made explicit.

Non-coherent accumulations are the sum of the squares of M coherent accumulations.

$$P(\hat{S}_k) = \sum_{m=1}^{M} I_m(\hat{S}_k)^2 + Q_m(\hat{S}_k)^2 \tag{3.3}$$

Non-coherent accumulations are referred to as "non-coherent" because carrier phase information is destroyed by the squaring operation prior to the final sum in Eq. (3.3). As a result, carrier phase need not be continuous at the boundaries of coherent accumulation intervals. This situation is illustrated in Figure 2.2. Allowing occasional phase discontinuities allows the total accumulation time to be lengthened beyond what would be possible without discontinuities.

The key to coarse acquisition will be to seek the $\hat{S}_k$ that maximizes $P(\hat{S}_k)$. Maximizing $P(\hat{S}_k)$ is equivalent to minimizing the sum of the squared fit errors of the signal model, Eq. (2.22). After optimizing X_m and Y_m, the sum of the squared errors in the signal model over the m^{th} coherent accumulation interval is

$$J = \frac{1}{2} \sum_{i=1}^{N} \left(y_{Ii}^2 + y_{Qi}^2 \right) - \frac{1}{2N} \left(I_m(\hat{S}_k)^2 + Q_m(\hat{S}_k)^2 \right). \tag{3.4}$$

Given that the first term is independent of $\hat{S}_k$, maximization of $P(\hat{S}_k)$ amounts to minimization of the sums of the squared errors over the M coherent accumulation intervals.

3.2 Discrete Fourier Transform-Based Accumulations

This section describes an algorithmic change that can be employed to speed the signal detection calculations. It may be much more efficient to compute the accumulations in the frequency domain than it is to compute them directly in the time domain, as in Eqs. (3.1) and (3.2). This is true if one must calculate many accumulations for many different offsets of the subcarrier with respect to the carrier, *i.e.*, many different values of τ_{B0}. These offsets are buried in $\boldsymbol{a}$ and $\boldsymbol{b}$ above. As it turns out, for the experiment described below, the transmission time of the code (that is, the data-laden subcarrier) was uncertain and necessitated searching values of τ_{B0} which varied up to a few seconds. In this case, it was several thousand times cheaper to calculate the accumulations in the frequency domain than in the time domain. This is important, because even the more-efficient frequency-domain calculations took several seconds to several minutes to complete. Section 3.6.4 discusses the computational burden of the acquisition algorithms.

The FFT-based algorithm is adapted from FFT-based acquisition algorithms employed in GPS processing [8]. The speed gains are due to the remarkable efficiency of the fast Fourier transform (FFT) algorithm employed to transform between the time and frequency domains for purposes of calculating the cross-correlation between two signals at many different offsets.

The process of calculating a discrete Fourier transform-based detection statistic involves several steps. The first step is to wipe off the carrier. The next step is to calculate the discrete Fourier transform (DFT) of the base-band mixed data samples and optionally filter them in the frequency domain. This leaves only the relatively low-bandwidth subcarrier. The final few steps are to generate the subcarrier and data bits replica using the candidate $\hat{S}_k$, take its DFT, calculate its complex conjugate, then multiply that together with the base-band mixed and DFTed samples, element-by-element. This sub-carrier replica functions much like a PRN code in standard GPS signal acquisition, except that it is complex-valued and can

take on values other than +1 and -1. Taking the inverse DFT of the result produces the correlations, or accumulations, at all code offsets considered.

Reviewing Eq. (2.17), it can be readily observed that baseband mixing to wipe off the carrier can be accomplished by multiplying each complex sample by the negative complex exponential of the estimated carrier phase. The result is the baseband mixed sample.

$$y_{BB}(\hat{S}_k, i) = y_i \exp\left(-j\alpha_i(\hat{S}_k)\right) \tag{3.5}$$

Figure 3.1 shows four views of the real and imaginary parts of the baseband mixed signal versus time for an on-orbit satellite. The i^{th} circle in the figure corresponds to $y_{BB}(\hat{S}_k, i)$ for the data set and parameter set used to generate the figure. The lower left panel of Fig. 3.1 clearly shows how the phase modulation of the carrier due to the subcarrier causes the phase of the carrier to vary from the nominal, unmodulated phase by about ± 1 radian. In this example, there are two subcarriers active, so the sinusoidal variation with regular 180° phase shifts is obscured by the addition of another sinusoid at a different frequency.

Define the vector of base-band mixed measurements

$$\boldsymbol{y}_{BB}(\hat{S}_k) = \left[y_{BB}(\hat{S}_k, 1),\ y_{BB}(\hat{S}_k, 2),\ \ldots,\ y_{BB}(\hat{S}_k, i),\ \ldots,\ y_{BB}(\hat{S}_k, N_{batch})\right]^{\mathrm{T}} \tag{3.6}$$

The indices $1 \ldots N_{batch}$ as given are for the first coherent accumulation interval in the first non-coherent accumulation. The indices must be adjusted for other intervals. N_{batch} must be chosen to be as large as necessary to contain the N samples over which accumulation is to be performed and an additional amount of samples that span the time over which the code start time is to be searched. Each coherent accumulation advances the FFT data batch indices by N samples. For instance, the second coherent accumulation would use samples $N + 1 \ldots N + N_{batch}$ and the third would use samples $2N + 1 \ldots 2N + N_{batch}$ and so on. In order to fully exploit the computational efficiency of the FFT algorithm, N_{batch} should be

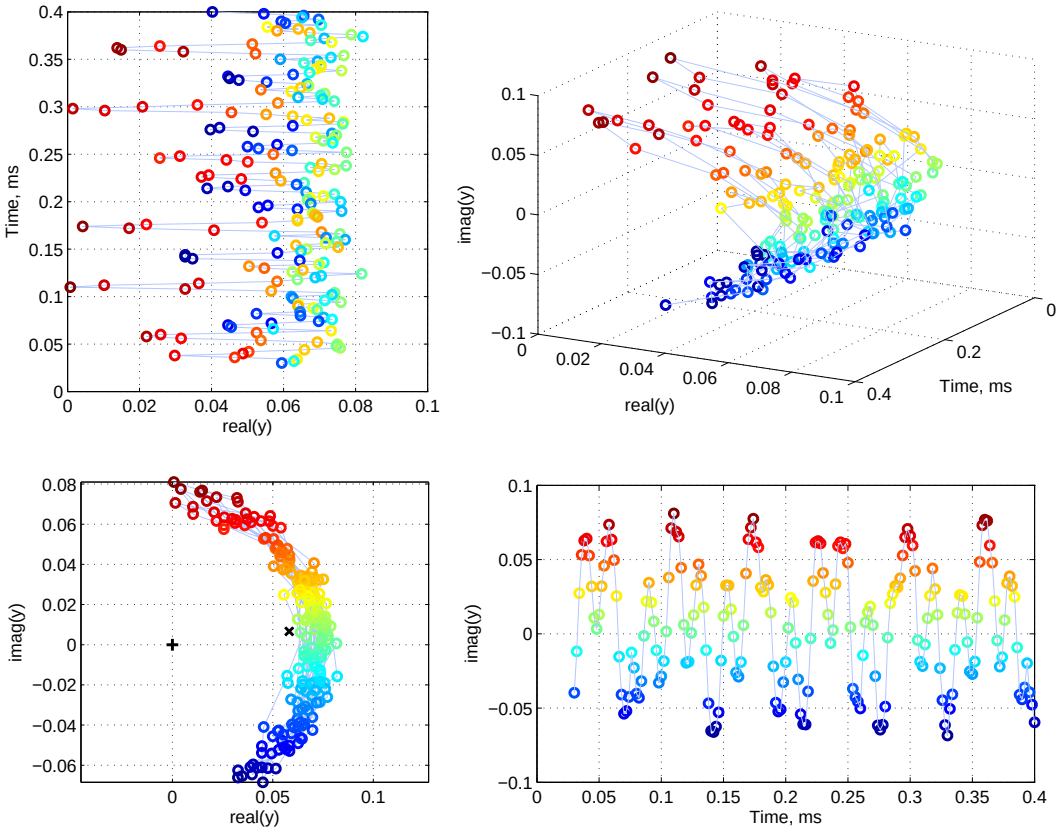


Figure 3.1 – The figures show the relationships of the real and imaginary parts of the baseband-mixed samples and time. Each sample is shown as a colored circle, with color representing the value of the imaginary part of the sample. The black 'x' in the lower left figure represents the average value of the samples. The samples are from an on-orbit satellite and were captured using the 13-m parabolic antenna. The remaining phase modulation of the signal after baseband mixing manifests as oscillations through a short arc along the surface of a cylinder in real(y), imag(y), t space.

chosen to be a power of two. The discrete Fourier transform of $\boldsymbol{y}_{BB}(\hat{S}_k)$ is $\boldsymbol{Y}_{BB}(\hat{S}_k)$.

$$\boldsymbol{Y}_{BB}(\hat{S}_k) = \text{FFT}\left(\boldsymbol{y}_{BB}(\hat{S}_k)\right) \tag{3.7}$$

After baseband mixing, it can be advantageous to filter and downsample the signal to the relatively low sample rate necessary to adequately sample the subcarrier signal. Applying a frequency domain filter, $\boldsymbol{H}(f)$, is a matter of multiplying two vectors together, element-by-element, $Y_{BB}(\hat{S}_k, f)H(f)$, at the frequencies f for which the FFT in $Y_{BB}(\hat{S}_k, f)$ has

been computed. Care must be taken in this step to match the correct elements of $\boldsymbol{Y}_{BB}(\hat{S}_k)$ with $\boldsymbol{H}(f)$. Most implementations of the discrete Fourier transform return the frequency components in a wrapped order. That is, from zero to increasingly high frequency, then most negative frequency to increasingly less negative. The base-band mixed signal must be filtered in such a way as to preserve all the signal power and reject as much noise power as possible. The present work uses a root raised cosine filter

$$H(f) = \begin{cases} 1 & |f| < \frac{1-\beta}{2T} \\ \sqrt{\frac{1}{2}\left(1 + \cos\left[\frac{\pi T}{\beta}\left(|f| - \frac{1-\beta}{2T}\right)\right]\right)} & |f| \geq \frac{1-\beta}{2T} \text{ and } |f| \leq \frac{1+\beta}{2T} \\ 0 & |f| > \frac{1+\beta}{2T} \end{cases} \tag{3.8}$$

with parameters $\beta = 0.3$ and $T = (1-\beta)/(2 \cdot 96000)$. The standard parameter name "β" unfortunately duplicates the name given to the subcarrier replica phase, but it appears as a filter parameter only this once. The parameters chosen produce a filter that does not significantly attenuate the telemetry signal, whose frequency response falls off fairly rapidly outside of the pass band, and whose time-domain side lobes are relatively short-lived. A plot of the power spectra of the filter and simulated telemetry signal are shown in Figure 3.2. As can be seen from the figure, most of the power of the telemetry signal is located inside the pass band of the filter.

After application of the frequency-domain filter, many high-frequency FFT bins contain no power. The excess high-frequency bins, those bins associated with frequencies above the cutoff frequency $(1+\beta)/(2T)$, can be deleted. Call the resulting vector $\tilde{\boldsymbol{Y}}_{BB}(\hat{S}_k)$. This vector is a different length from $\boldsymbol{Y}_{BB}(\hat{S}_k)$. The inverse FFT of $\tilde{\boldsymbol{Y}}_{BB}(\hat{S}_k)$, say $\tilde{\boldsymbol{y}}_{BB}(\hat{S}_k)$, is effectively a filtered and downsampled version of $\boldsymbol{y}_{BB}(\hat{S}_k)$ covering the same time span.

The choice of cutoff frequency affects the final performance of the FFT-based grid search algorithm. In order to keep the FFT and IFFT as efficient as possible, the number of remaining frequency bins should be chosen to be a power of two. If the cutoff frequency of

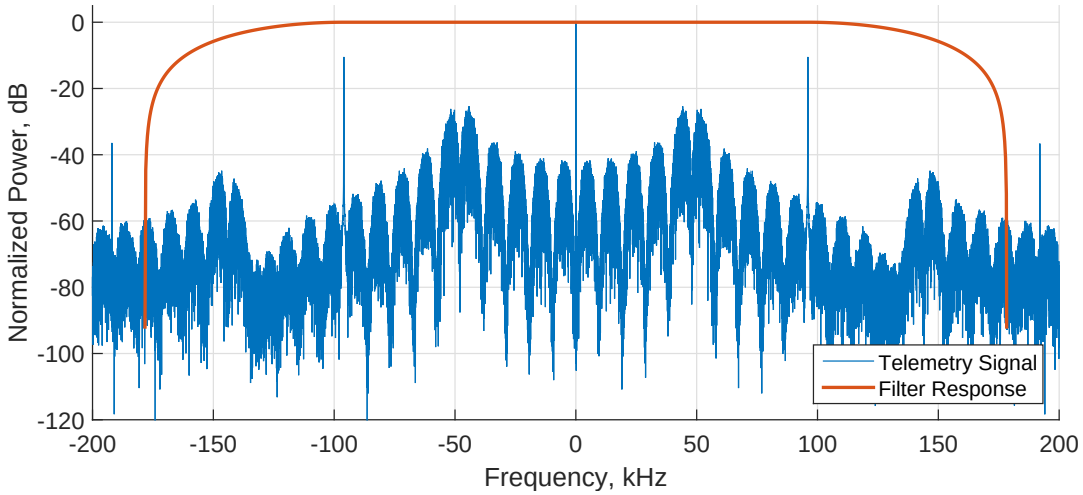


Figure 3.2 – The power transfer of the filter (red, thick) is plotted along with the normalized power spectrum of the telemetry signal (blue). The pass band of the filter has unity gain and passes the majority of the signal power.

the filter is $f_0 = (1+\beta)/(2T)$, then the number of frequency bins inside, or partially inside, the range 0 to f_0 is $N_{\Delta f} = \text{ceil}\,(f_0/\Delta f)$, where the function "ceil" rounds its argument to the next larger integer, and the frequency bin width in $\boldsymbol{Y}_{BB}(\hat{S}_k)$ is $\Delta f = f_{samp}/N_{batch}$. The power of two we want is then $p = \text{ceil}(\log(2N_{\Delta f})/\log(2))$. The factor of 2 inside the logarithm is included because both positive and negative frequencies must be accounted for. The new effective sample rate is $f_{sub} = 2^p \Delta f$. The new sample rate will be a power-of-two fraction of the original sample rate and the number of samples in $\tilde{\boldsymbol{y}}_{BB}(\hat{S}_k)$ will be $N_{sub} = N_{batch} f_{sub}/f_{samp}$.

This optional filtering step is the final operation to be performed on the base-band-mixed signal alone. The remainder of the operations prior to computing the actual accumulations are applied to a subcarrier and data bits replica signal. The subcarrier replica must be generated at the same sample rate f_{sub} as $\tilde{\boldsymbol{y}}_{BB}(\hat{S}_k)$. However, it will not be the same length. The subcarrier replica will consist of as many samples as are required to span the target coherent accumulation interval at the sample rate f_{sub}. The number of subcarrier samples will be $N_{code} = \text{round}\,(N f_{sub}/f_{samp})$, where the function "round" rounds its argument to the

nearest integer. The rounding step is necessary because the product of N and f_{sub}/f_{samp} need not be a whole number (*e.g.* for $N = 1050$ and $f_{sub}/f_{samp} = 1/4$). If it is not a whole number, the subcarrier replica length will vary slightly from the desired length.

The l^{th} subcarrier replica sample is $x_{code}(\hat{S}_k, l) = \exp\left(j\beta_l(\hat{S}_k)\right)$, for $l = 1...N_{code}$, or simply the vector of all N_{code} samples $\boldsymbol{x}_{code}(\hat{S}_k) = \left[x_{code}(\hat{S}_k, 1), x_{code}(\hat{S}_k, 2), ..., x_{code}(\hat{S}_k, N_{code})\right]^{\text{T}}$. The subscripts have changed from i to l, because if the baseband mixed signal was filtered and downsampled, these samples are no longer associated with the same sample times as those which applied to the signal samples and the carrier replica samples.

The subcarrier replica must be zero-padded to contain the same number of samples as frequency bins in $\tilde{\boldsymbol{Y}}_{BB}(\hat{S}_k)$. The zero padding causes the FFT of the subcarrier vector to have the same frequency resolution as $\tilde{\boldsymbol{Y}}_{BB}(\hat{S}_k)$. The zero-padded subcarrier vector is constructed as $\boldsymbol{x}(\hat{S}_k) = \left[\boldsymbol{x}_{code}(\hat{S}_k)^{\text{T}}\ 0\ \ldots\ 0\right]^{\text{T}}$ with $N_{sub} - N_{code}$ zeros. The FFT of this padded subcarrier vector is $\boldsymbol{X}(\hat{S}_k) = \text{FFT}(\boldsymbol{x}(\hat{S}_k))$.

The final step in calculating the accumulations at multiple code offsets is to take the inverse FFT of the complex conjugate of the padded subcarrier FFT multiplied by the FFT of the filtered data. Let

$$\boldsymbol{Z}(\hat{S}_k) = \begin{bmatrix} \tilde{Y}_{BB}(\hat{S}_k, 1)X(\hat{S}_k, 1)^* \\ \tilde{Y}_{BB}(\hat{S}_k, 2)X(\hat{S}_k, 2)^* \\ \vdots \\ \tilde{Y}_{BB}(\hat{S}_k, i)X(\hat{S}_k, i)^* \\ \vdots \\ \tilde{Y}_{BB}(\hat{S}_k, N_{sub})X(\hat{S}_k, N_{sub})^* \end{bmatrix} \tag{3.9}$$

where the star, '*', denotes the complex conjugate. The complex conjugate is taken in order to mix off the subcarrier and leave only a DC value. This is equivalent to multiplying by the negative complex exponential of subcarrier phase. The accumulations in the time domain

are simply

$$\boldsymbol{z}(\hat{S}_k) = \text{IFFT}\left(\boldsymbol{Z}(\hat{S}_k)\right) \tag{3.10}$$

The vector $\boldsymbol{z}(\hat{S}_k)$ consists of the accumulations at all code offsets from 0 to $N_{sub} - 1$. Thus, the output of the IFFT operation in Eq. 3.10 computes accumulations for a range of τ_{B0} signal parameter values all at once, thereby facilitating a rapid signal acquisition search for the value of this unknown quantity. Each code offset will consist of a delay of $1/f_{sub}$ seconds from the initial value specified in $\hat{S}_k$. The algorithm will compute N_{sub} code offsets; however the last $N_{code}-1$ offsets will be meaningless due to the violation of the periodicity assumption of the FFT. Only the first $N_{sub} - N_{code} + 1$ code offsets will be meaningful.

Additionally, the non-causal nature of the frequency-domain filter will introduce some dependency of each time-domain sample on several surrounding samples—both before and after a given sample—due to the corresponding time-domain impulse response function of the $H(f)$ filter. For the filter described above, samples within about 5×10^{-5} s on either side of the current sample contribute significantly to the filtered value. So, the first and last several code offsets must also be discarded, because the value of the filtered data vector at those samples is based partially on values from the other end of the unfiltered sample vector. Figure 3.3 illustrates the various ranges of valid code offset.

If no filter were applied, accumulations computed in this way would have exactly the same value as those computed by brute force via Eqs. (3.1) and (3.2). If a filter is applied, the two methods yield only approximately the same results. That is,

$$z_m(\hat{S}_k, l) \approx I_m(\tilde{S}_k) + jQ_m(\tilde{S}_k) \tag{3.11}$$

where the l subscript in the argument of $z_m(\hat{S}_k, l)$ is the one that corresponds to the same value of τ_{B0} used in Eqs. (3.1) and (3.2), $\tau_{B0} = (l-1)/f_{sub} + t_k$, and the circumflex over the parameter set on the right-hand side of the equation has been replaced by a tilde to indicate that the value of τ_{B0} in $\tilde{S}_k$ may not be the same as in $\hat{S}_k$. These accumulations take

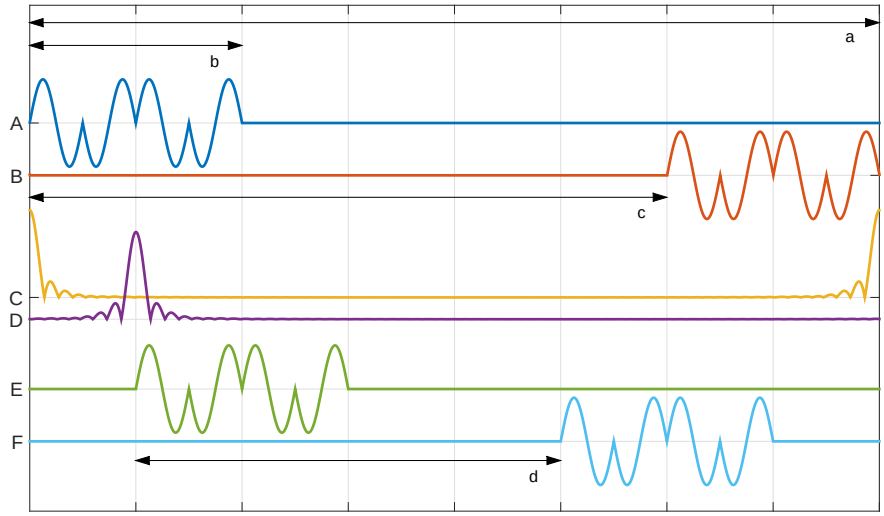


Figure 3.3 – This figure illustrates time domain relationships between the data interval, the code replica and its offsets, and the time-domain impulse response of the filter applied in the frequency domain. Double arrow a shows the data interval of N_{sub} samples. Double arrow b shows the length of the subcarrier replica N_{code} samples. Double arrows c and d show the range of valid code offsets for the unfiltered and the filtered data, respectively. Trace A is a cartoon representation of the subcarrier replica padded with trailing zeros. Note that the number of subcarrier cycles per Manchester chip has been reduced to one for illustrative purposes. Trace B is the same subcarrier replica padded with trailing zeros, but at a code delay of $N_{sub} - N_{code}$ samples. Traces C and D show the magnitude of the filter impulse response, both at a delay of 0 samples and at the first useable delay – the one at which the total energy under the impulse response is negligible. Traces E and F show the zero-padded subcarrier replica shifted to the first and last useable code delays when the filter has been applied to the data.

thousands of times less computation time for coherent accumulations in the tens of seconds and searched over approximately 2 seconds of code start time offset, τ_{B0}.

In order to compute non-coherent accumulations from these FFT-based coherent accumulations, one need only sum the squares of the magnitudes of $z_m(\hat{S}_k, l)$ over all coherent intervals, m, and for each code offset, l. Explicitly,

$$P(\hat{S}_k, l) = \sum_{m=1}^{M} \| z_m(\hat{S}_k, l) \|^2 \tag{3.12}$$

3.3 Hypothesis Test

This section is largely a review of standard signal detection theory, but with notation consistent with the foregoing content of the dissertation. It is the hypothesis test that provides a formalism for determining whether or not a signal has been detected. In order to establish whether the non-coherent accumulation of Eq. (3.3) or (3.12) represents the accumulation of only noise or of signal and noise, the value $P(\hat{S}_k)$ (or $P(\hat{S}_k, l)$, a distinction which will be dropped from here forward) must be compared to a threshold γ_F. After the comparison, we either accept hypothesis H_0, that the signal is not present, or H_1, that the signal is present. Under H_0, with probability of false alarm α_F, $P(\hat{S}_k) \leq \gamma_F$. Under H_1, with probability of detection α_D, $P(\hat{S}_k) > \gamma_F$.

The threshold γ_F can be derived by assuming that under the null hypothesis, H_0, each $I_m(\hat{S}_k)$ and $Q_m(\hat{S}_k)$ for $1 \leq m \leq M$ is drawn from a normal distribution with zero mean and unity variance. The condition on variance can be enforced via normalization by simply dividing the accumulations by their standard deviation such that the resulting normalized accumulations have unity variance. In this case, $P(\hat{S}_k)$ will be the sum of the squares of $2M$ values drawn from a normal distribution with zero mean and unity variance, that is $P(\hat{S}_k)$ will be drawn from a χ^2 distribution with $2M$ degrees of freedom [6].

$$p(\gamma|H_0) = \frac{1}{2^M(M-1)!}\gamma^{M-1}e^{-\gamma/2}; \gamma \geq 0 \tag{3.13}$$

Generally, α_F, the probability of false alarm, is chosen to equal a desired value, and γ_F is determined in order to achieve this value. This value applies not to any individual instance of the signal detection statistic, but to the entire batch of, say, N_s statistically independent detection statistics calculated in a grid search. This is because the χ^2 distribution has support over the entire positive real axis, so if enough accumulations of pure noise are performed, eventually any given detection statistic can be attained. The threshold value, γ_F, is chosen such that it is unlikely that for the number of independent accumulations performed, the

statistic will exceed the threshold if no signal is present. This is achieved by choosing γ_F to satisfy

$$(1 - \alpha_F)^{1/N_s} = \int_0^{\gamma_F} p(\gamma|H_0)d\gamma \tag{3.14}$$

This calculation can be carried out in MATLAB, for instance, with the `chi2inv` function.

Under H_1, $I_m(\hat{S}_k)$ and $Q_m(\hat{S}_k)$ will be drawn from a normal distribution with a non-zero mean, but with the same variance as under H_0. In this case, the resulting distribution of $P(\hat{S}_k)$ will be a non-central χ^2 distribution. This distribution will have $2M$ degrees of freedom and non-centrality parameter $2M\beta$.

$$p(\gamma|H_1) = \frac{1}{2}\left(\frac{\gamma}{2M\beta}\right)^{\frac{1}{2}(M-1)} e^{-\frac{1}{2}(\gamma+2M\beta)} \mathrm{I}_{M-1}\left(\sqrt{2M\beta\gamma}\right); \gamma \geq 0 \tag{3.15}$$

Above, $\beta = (C/N_0)N/f_{samp}$ is the two-sided pre-detection carrier power-to-noise power ratio and $\mathrm{I}_{M-1}(\sqrt{2M\beta\gamma})$ is a modified Bessel function of the first kind [6]. The name β unfortunately duplicates earlier nomenclature, however the current β only appears within this section. With the threshold value chosen based on the desired probability of false alarm, the probability of detection, α_D, is determined from

$$\alpha_D = 1 - \int_0^{\gamma_F} p(\gamma|H_1)d\gamma \tag{3.16}$$

That is, the γ_F that corresponds to the chosen probability of false alarm, the actual C/N_0, the coherent integration time, and the number of coherent integrations determine the probability of detection. For each point in the signal parameter space grid search, if the signal is present with those parameters, it will be detected with probability α_D. If no signal is present, of all N_s statistically independent points in the grid search, one of them may exceed the detection threshold γ_F and result in a false alarm with probability α_F.

This is, in effect, an *ad hoc* M-ary hypothesis test [9, 10] with one hypothesis for each point in the parameter space grid and one hypothesis that no signal is present. In practice, if a signal is present in the samples, there will be a volume in the signal parameter space

which produces detection test statistics which exceed the threshold level. For this *ad hoc* test, the point which produces the highest $P(\hat{S}_k)$ which exceeds the threshold γ_F should be chosen as the true hypothesis. If no point produces a $P(\hat{S}_k)$ which exceeds γ_F, then the null hypothesis should be accepted.

Since the probability of detection depends on parameters chosen by the system engineer, it can be adjusted. In order to increase the probability of detection, the separation of the two distributions must be increased. The means of the two distributions provide part of the picture with regard to their overlap. In general the farther apart their means, the less overlap they will share. In order to increase the mean of the signal-present PDF, the non-centrality parameter $2M\beta$ must be increased. There are two ways of increasing the non-centrality parameter for a given signal and sampling rate. Either the number of samples N in each coherent accumulation can be increased, or the number of coherent accumulations M per non-coherent accumulation can be increased. Increasing N is desirable, because it raises the mean of only the non-central χ^2 signal-present distribution. Increasing M increases the means of both the χ^2 no-signal distribution and the non-central χ^2 signal-present distribution. The length of the coherent accumulations is limited by the stability of both the carrier-generating spacecraft clock and the receiver clock. As the coherent accumulation interval increases, random walk-type processes, *e.g.* flicker phase noise, become more important, as these processes do not average out [11]. These random processes are not modeled well by the quadratic carrier phase approximation in Eq. (2.10). This poor modeling results in effective power loss, which translates into a lower β value in the above analysis, thereby reducing the detection power instead of increasing it.

Each detection statistic calculated uses a unique set of signal parameters $\hat{S}_k$. In order to be certain that the signal is detected if it is present, the detection statistic must be calculated using a set of signal parameters that is sufficiently close to the true or optimal values. In practice, this means calculating detection statistics over a multi-dimensional grid of signal

parameter values. The choice of grid is the topic of the next section.

3.4 Grid Search

When performing a signal detection, the signal parameter space must be explored. The method used in this work is a three-dimensional grid search: the dimensions searched are carrier Doppler shift ω_{DAk}, subcarrier Doppler shift ω_{DB0}, and code start time τ_{B0}. As was discussed in Section 3.2, the code start time dimension is searched over the grid set by the FFT-based grid search algorithm. The two parameters not explicitly considered during the grid search, the two Doppler rates, are optimized during the fine acquisition phase described in Chapter 4. These parameters are not searched during the coarse acquisition, because over short intervals the rate of change of Doppler shift is small and making the approximation that the Doppler is constant results in only a small power loss. For instance, over a 10-second accumulation, a carrier Doppler rate of 5×10^{-3} Hz/s results in an accumulated phase error of only 0.1 cycles and 93.5% of the power is recovered. The values of the Doppler rates encountered during the proof of concept experiment are significantly smaller than 5×10^{-3} Hz/s. In a situation in which higher rates are expected, an *a priori* estimate of Doppler rate could be used to reduce power loss but keep the search space tractable.

The grid spacing in each of the Doppler dimensions must be sufficiently tight not to miss any important features, but simultaneously as loose as possible in order to reduce computational demand. This section of the report deals with the choice of grid for these two search dimensions. Both the absolute range of the grid, and the spacing within that range, must be determined for each dimension.

In order to determine the grid used for carrier Doppler, we need to consider the accuracy of spacecraft clock A and its frequency stability. The former affects the range of the grid,

and the latter affects the grid spacing through the choice of coherent accumulation interval. The coherent accumulation interval should be chosen such that the clock rate drift over the interval causes less than about one quarter of a cycle of phase error. A worst-case short coherent accumulation interval can be chosen based on the clock stability specification levied on the spacecraft oscillators. The specified Allen deviation times the nominal carrier frequency is less than 3 Hz for gate times from 0.1 to 1 seconds. In order to accumulate no more than about 0.25 cycles of phase error on average, the accumulation interval must be no longer than 0.25 cycles/3 Hz = 0.0833 s. In practice, much longer accumulations can be used, because the oscillators perform better than specified. For instance, for a data set gathered in the factory at SSL, the clock stability was sufficient to integrate for about 0.4 seconds before starting to lose power. For other data sets taken from an on-orbit satellite, the coherent accumulation interval could be as long as about 0.25 seconds.

The length of the coherent accumulation interval affects the required grid spacing. The longer the coherent accumulation, the tighter the required grid spacing, because accumulations will be more sensitive to errors in carrier frequency. In order to retain significant power, the grid spacing should be set such that the accumulated phase error over a coherent accumulation interval due to mismatch between the grid frequency and the true average frequency does not exceed about 0.25 cycles. If the grid spacing is, say, S_p Hz, then the maximum carrier frequency error is $S_p/2$ Hz. S_p must be set to satisfy $\frac{S_p}{2} T_{coh} \leq 1/4$ cycles, or $S_p \leq \frac{1}{2T_{coh}}$ Hz. For instance, for 0.125-s accumulations, the carrier Doppler grid spacing must be no greater than 4 Hz and for 0.25-s accumulations, the spacing must be no greater than 2 Hz.

The absolute range of the carrier Doppler grid is determined by the accuracy of spacecraft clock A. The specified accuracy is ±1 PPM, which translates to ±12.6985 kHz for a nominal carrier frequency of 12.6985 GHz—the carrier frequency of the test satellite. This range is enormous when gridded at 4 Hz intervals and even larger when gridded at the finer

resolutions required for longer coherent accumulations. However, if an estimate of the carrier Doppler shift is available, it can be used to narrow the search space. In the proof-of-concept experiment, such an estimate was available from samples taken simultaneously using a very high-gain antenna which resulted in a high signal power-to-noise power ratio (SNR). During operation of a deployed orbit-determination system, such an estimate would be provided as part of the handoff from the traditional orbit determination system in which a high-gain antenna is employed to perform satellite ranging. After handoff, the proposed low-cost orbit determination system would measure the spacecraft clock frequency error as part of its estimation process. Each measurement performed by the network of ground stations would be informed by the past history of measurements. The spacecraft clock frequency error estimate and its variance would be propagated forward in time along with an estimate of the expected Doppler shift due to satellite motion. This information would limit the Doppler frequency search range for each measurement to a range much smaller than the bounds implied by the spacecraft clock's specification for maximum frequency error.

As for carrier Doppler, the grid spacing for subcarrier Doppler must be calculated based on a maximum tolerable phase error over the coherent interval. For the subcarrier, the coherent interval is the entire non-coherent accumulation interval. In general, if there are M coherent accumulations each of T_{coh} s duration, then the subcarrier Doppler grid spacing should be set equal to or less than $\frac{1}{2MT_{coh}}$ Hz. For instance, if $T_{coh} = 0.125$ s and $M = 80$, then the subcarrier Doppler grid spacing should be less than or equal to 0.05 Hz. Or, if $T_{coh} = 0.125$ s and $M = 12$, then the subcarrier Doppler grid spacing should be less than or equal to 0.33 Hz .

Again like carrier Doppler, the range of the subcarrier Doppler grid is set by the accuracy of the spacecraft clocks. In this case, it is spacecraft clock B which is important. Its accuracy is specified to be within ± 50 PPM. For a nominal subcarrier frequency of 48 kHz, this means there could be as much as 2.4 Hz of subcarrier Doppler. This range may be reduced if

an estimate of subcarrier Doppler is available. For the proof-of-concept experiment, this estimate came from the high-SNR signal path. For a deployed orbit determination system, the dynamical model of the spacecraft could include a model of the subcarrier clock error dynamics which would provide this estimate. The dynamic model presented in Chapter 5 does not currently include such an estimate.

With the grid spacing and grid range determined, the grid search consists of computing the detection test statistic at each grid point. The detection test statistic is computed as described in the preceding sections, with $\hat{S}_k$ set for each grid point. During the proof-of-concept experiment, the extent of the grid search was minimal due to the exceptionally clean estimates of carrier- and code-Doppler-shift available from the high-SNR signal path.

3.5 Proof-of-Concept Experiment

On 11–13 March 2014, at the Echostar Satellite Services facility in Gilbert, Arizona, the feasibility of the low-cost orbit determination concept was evaluated in a proof-of-concept experiment. The experiment consisted of a data gathering phase and an analysis phase. Data gathering and preliminary analysis occurred on site. The final analysis was conducted after the field campaign ended. Prior to the end of the field campaign, it was demonstrated that the RF and digitization equipment was functioning as expected and that the spacecraft telemetry signal had been positively detected on the low-SNR signal path using a standard gain horn (SGH) antenna. Acquisition of the signal in samples from an omnidirectional antenna was also demonstrated after returning from the field.

The remainder of this section describes the experimental setup and procedures. The next sections will present some results of the analysis of the data from the experiment.

3.5.1 Experimental Setup

The hardware for the experiment was configured as shown in Figure 3.4. In the block diagram, there are two signal paths, the high-SNR and the low-SNR paths. The high-SNR path was recorded in order to provide a "truth" against which to compare the results from the low-SNR path. The high-SNR path starts at the 13-m antenna and goes through the high-power amplifier (HPA) building, where down conversion to a 70 MHz intermediate frequency (IF) occurs. The down converter was supplied with a 10-MHz reference signal. The 70-MHz signal then passed out of the building and was carried to the experimental setup in a coaxial cable. The signal was band-pass filtered prior to being digitized on the first channel of the dual-channel universal software radio peripheral (USRP). The 70-MHz signal was also split out to a Cortex modem, not shown in the block diagram, in order to record the data bits as transmitted by the satellite.

The low-SNR path starts at either the 15-cm standard gain horn (SGH), shown in the block diagram and in Fig. 3.5, or at an omnidirectional antenna shown in Fig. 3.6 which replaced the SGH in some data collection runs. The signal from that antenna is amplified with a low-noise amplifier, then mixed down to positive 70 MHz IF using a mixing signal generated by a direct digital synthesis frequency synthesizer which was provided with the same 10-MHz reference signal used in the down conversion of the high-SNR signal path. After mixing, the signal passed through a 1 GHz low-pass filter (LPF) to kill any local oscillator (LO) power leaking through the mixer. Finally, the signal was passed into the second channel of the USRP. The USRP also was supplied with the same 10-MHz reference signal used in both mixing chains. This allowed the USRP to mix both channels to exactly 0 Hz according to the 10-MHz reference. The USRP complex-sampled both channels simultaneously at 500 kilosamples per second. The digital samples were recorded by a computer connected to the dual-channel USRP via gigabit ethernet.

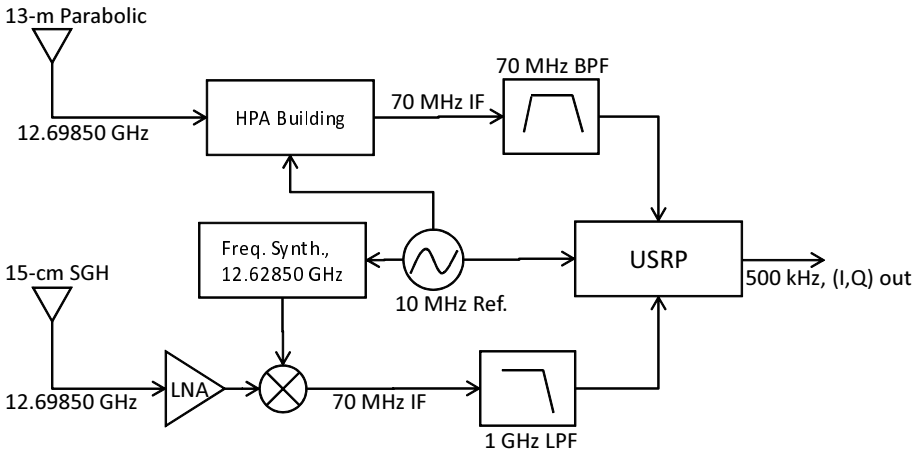


Figure 3.4 – Block diagram of experimental hardware.

The two antennas were pointed at the test satellite located at approximately -77°E longitude. The azimuth and elevation of the satellite from the earth station was calculated to be 128.4° azimuth by 35.8° elevation. The 13-meter parabolic antenna pointing was handled by Echostar. The 15-cm SGH was mounted to a tripod and pointed manually. The manual pointing procedure was *ad hoc*, but worked well. Elevation was measured using an analog inclinometer, Figure 3.7. Azimuth was measured using trigonometry and marks scratched on the ground with a rock relative to the North-South line established by the edge of the concrete antenna pad on which the experimental apparatus was positioned. A sketch of the geometry involved (not to scale) as well as an aerial view of the site is shown in Figure 3.8. In the figure, the experimental apparatus was located on the vacant antenna pad located immediately to the left of the overlaid triangle. The 13-m antenna used is the one furthest to the left.

3.5.2 Experimental Procedures

Prior to recording any data, the signals on the low- and high-SNR paths were verified to be present and reaching the USRP. This was accomplished by connecting the signal to a signal

Figure 3.5 – Photo of experiment hardware. The SGH is the pyramidal antenna on a tripod in the foreground. The actual 13-m antenna used for the experiment is in the background. This 13-m antenna is typical of those used for tone ranging and other two-way ranging techniques.

Figure 3.6 – Photo of experiment hardware. The omni-directional antenna used is sitting atop a tripod in the foreground of the image. In the background, is another 13-m antenna at the satellite uplink facility in Gilbert, AZ.

Figure 3.7 – Setting the elevation of the SGH antenna.

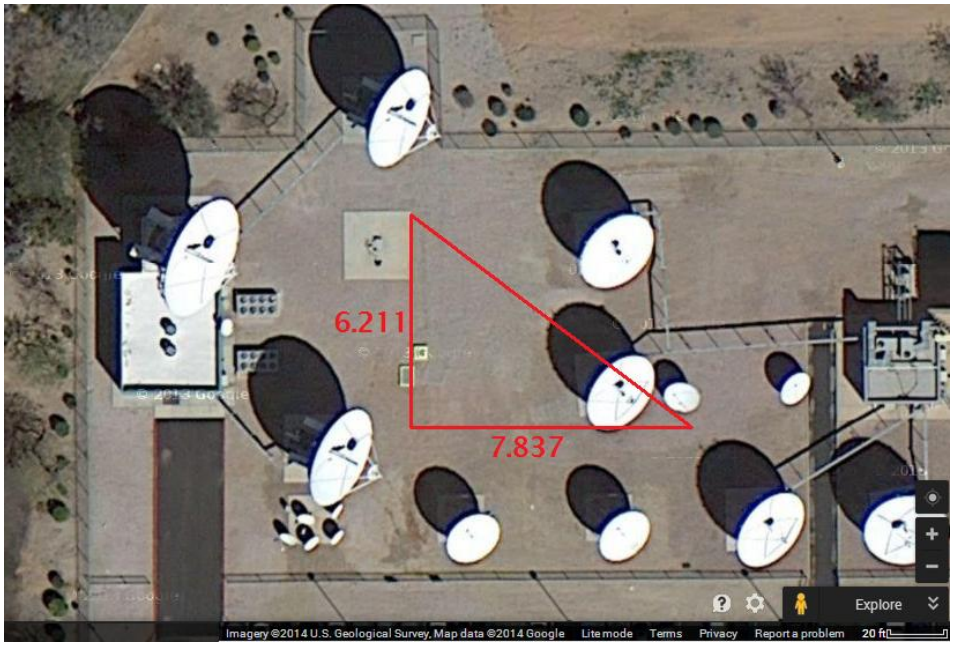


Figure 3.8 – Geometry used to align the antenna in azimuth (overlaid red triangle, not to scale), and an aerial view of the experiment site. The experimental apparatus was located on a table and equipment cart set up on the vacant antenna pad to the left of the apex of the overlaid triangle. The 13-m antenna due west of (*i.e.* to the left of) the vacant pad was the one used during the experiment.

analyzer rather than to the USRP. The signal was verified visually on the signal analyzer screen. For the high-SNR path, this was trivial. The entire signal sat well above the noise floor, Figure 3.9. C/N_0 was measured at 74.7 dB-Hz for the 13-m antenna. For the low-SNR path, checking the signal analyzer for evidence of the signal was possible for the SGH data but not for the omnidirectional antenna's data. The SGH provided just enough gain to raise the carrier peak above the noise floor. All of the power in the side bands was invisible below the noise floor, Figure 3.10. C/N_0 for the SGH was measured at 30.4 dB-Hz, which is not far from the 32 dB-Hz predicted from the link budget in Table 2.1. The carrier visible above the noise floor was at the same frequency as that visible on the high-SNR path[1]. The link budget predicts that the signal from the omni will be 21 dB down from the signal from the SGH. Since the signal fromthe SGH is only 13 dB up from the noise floor, it is clear that the signal from the omni is completely buried under the noise. But, since the only change between recording samples from the SGH and recording samples from the omni was the antenna itself, the experiment proceeded on the assumption that signal from the omni was reaching the USRP. C/N_0 is estimated to be about 11 dB-Hz for the omnidirectional antenna.

One batch of data was gathered on 11 March and five on 12 March. Each batch consisted of a set of samples recorded from the low-SNR path, a set of samples recorded from the high-SNR path, and a record of the data stream content recorded from the Cortex modem. The batch of data gathered on the 11^{th} used the SGH on the low-SNR path. Of the five batches recorded on 12 March, two of them used the SGH on the low-SNR path, and for the other three the SGH antenna was replaced by an omnidirectional antenna of the type used as the +Z omni for SSL spacecraft.

Each data collection run began by starting the Cortex modem recording the data stream. Five seconds after starting the Cortex, as measured by a digital stopwatch, the USRP was

[1]The two screen shots shown in Figures 3.9 and 3.10 were taken at different times, so the frequencies shown do not match.

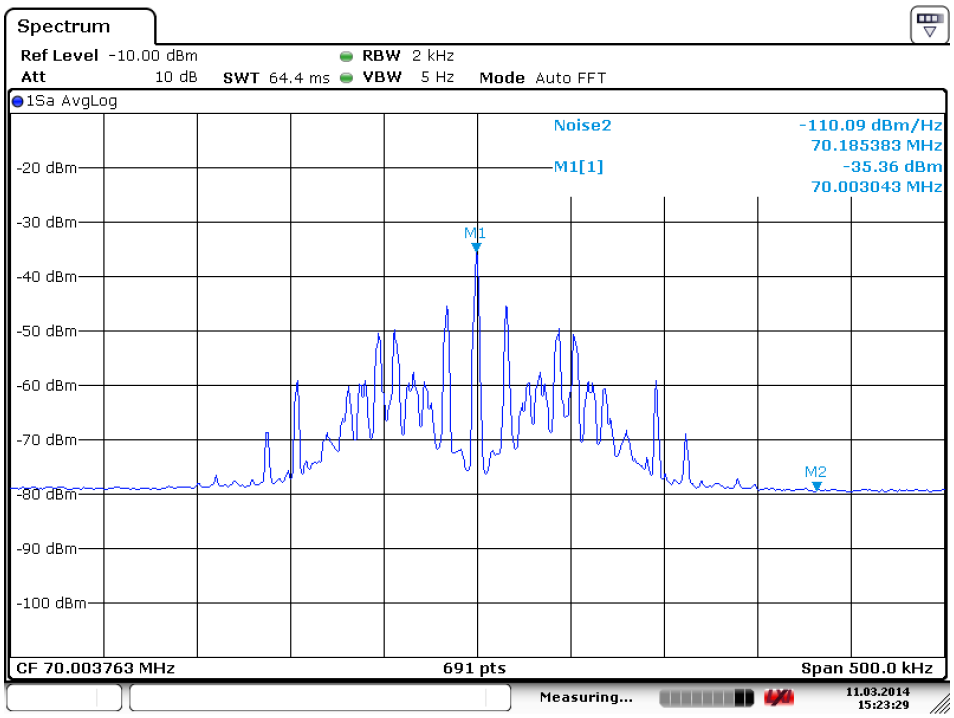


Figure 3.9 – Signal analyzer screen shot of high-SNR signal.

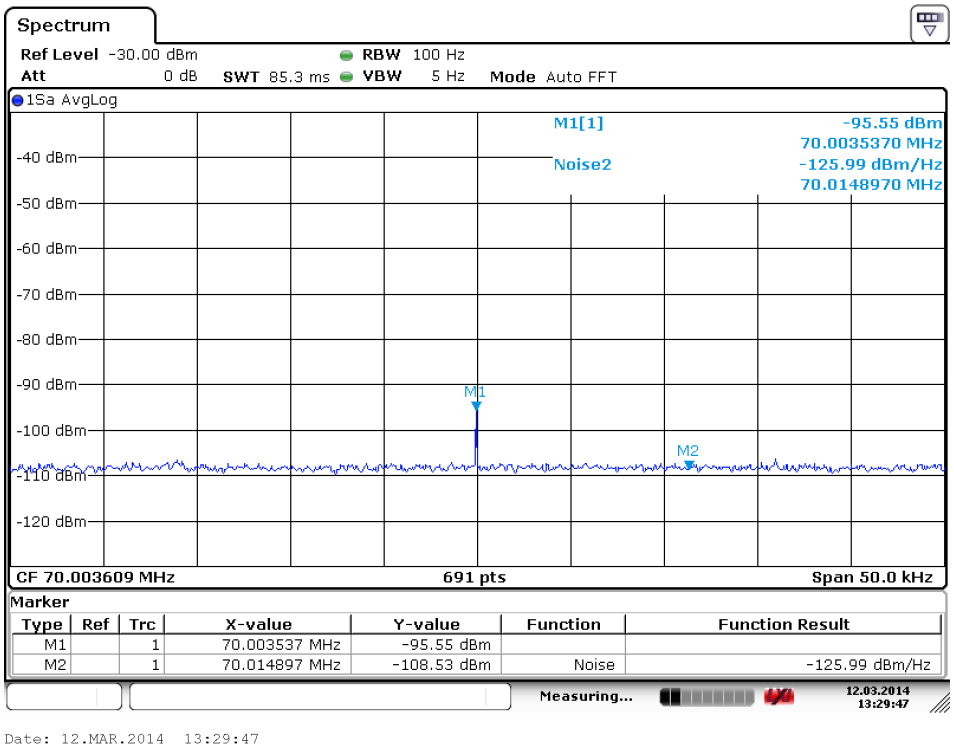


Figure 3.10 – Signal analyzer screen shot of low-SNR signal from SGH.

started. The Cortex recorded data for approximately 45 seconds or, for the runs using the omni antenna, 90 seconds. The USRP recorded until stopped manually. The USRP was typically stopped shortly before the Cortex finished recording.

3.6 Acquisition Results

The data gathered during the proof-of-concept experiment have been analyzed. One or more positive signal detections have been attained for each of the data sets investigated. The first step in each analysis is to establish the "truth" signal parameters from inspection of the high-SNR data. Next, the acquisition calculations are performed on the low-SNR data using Doppler frequency and rate estimates from the high-SNR data as a starting point. The results of each of these steps are discussed below.

3.6.1 Establishing the Truth

Using the high-SNR signal to establish the "true" carrier and subcarrier frequency history is fairly straightforward. For a given data set, the carrier frequency was found by picking the peak off of an FFT of 1-second batches of data. A polynomial fit through the peak frequencies was used as an initial guess for an interval-by-interval optimization of carrier frequency. The optimization was a Gauss-Newton optimization, similar to the one described in Section 4.1. Though, in this case, the optimization was for a single parameter, carrier Doppler frequency, one interval at a time. Interval duration was set to one 1/32 of a second. After finding the carrier frequency, the carrier was mixed off in order to simplify the search for subcarrier Doppler. Finding the second harmonic peak from the FFT of the baseband mixed samples gave a reasonable initial estimate of subcarrier Doppler. A 2-dimensional grid search was then carried out in which subcarrier Doppler and code start time were searched.

Following the grid search, the three subcarrier parameters were optimized in a procedure almost identical to the one in Section 4.1, except that this optimization was concerned only with the three subcarrier parameters.

There were some unexpected results. It was discovered that the signal frequency was undergoing step changes at intervals of whole seconds. Figure 3.11 shows carrier Doppler versus time for one batch of data. The frequency steps are clearly visible.

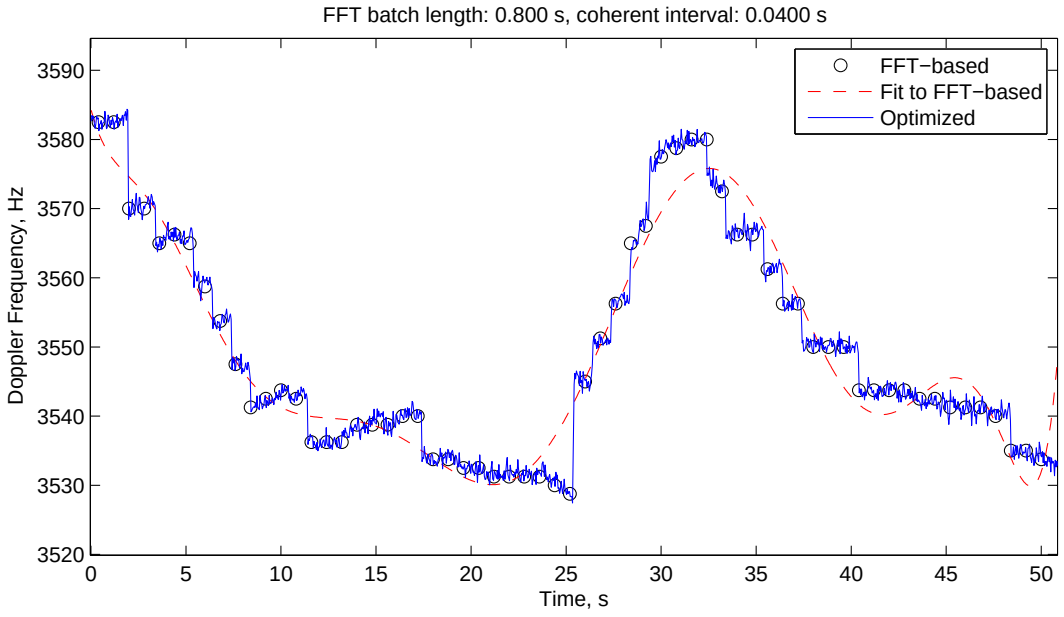


Figure 3.11 – Optimized carrier Doppler frequency (solid blue) is plotted over a 50-second interval. Also shown are the FFT-based frequency estimates (black 'o' markers) and the 10^{th}-order polynomial fit (dashed red) to the FFT-based estimates which was used as an initial guess for the optimization.

The origin of the frequency step changes is not known with certainty. They could originate either in the signal as transmitted by the satellite or in the mixing signal generated on the ground. No known source of such step changes in frequency exists on the satellite. The most probable cause of the step changes is one pulse-per-second (1 PPS) GPS discipline of the 10 MHz reference signal used in all the ground equipment. The GPS discipline is achieved by dithering the oscillator frequency around the nominal according to a feedback law derived from the timing of the 1 PPS signal. If the frequency jitter were a mere one part per billion,

or 0.01 Hz, this would result in changes of up to 12.6985 Hz when the nominally 10 MHz signal was multiplied up to 12.6985 GHz. The largest observed steps are between 10 and 20 Hz, which is consistent with a frequency jitter of about one part per billion due to 1 PPS GPS discipline.

The fact that the period of the steps and the amplitude of the steps is consistent with 1 PPS GPS discipline does not prove that that is the source; it merely points to a probable cause. An experiment would need to be carried out to determine with certainty the origin of the step changes in frequency. One simple experiment that could be performed would be to observe the signal on a signal analyzer with the 10 MHz reference successively active and inactive. If the carrier peak jumps left and right while the reference is active but not while it is inactive, this would be a strong indication that the 10 MHz reference was the culprit. Alternatively, a GPS receiver could be provided with the 10 MHz reference signal. If step changes in the receiver's clock rate estimate appear at regular 1-second intervals or multiples thereof, this is proof that the 10 MHz reference is the culprit.

In either case, if the step changes in frequency originate on the satellite or on the ground, there are some interesting lessons to be learned. If the spacecraft clock is experiencing step changes in oscillation frequency, then the signal processing algorithms must be designed to accommodate this behavior. It would be necessary to employ sufficiently short coherent intervals to prevent unacceptably large carrier phase errors from accumulating during coherent intervals. It would also be necessary to tie the samples at all ground stations to an absolute time reference so that samples containing the same features would be compared. Comparing one interval containing a step change to another not containing that same change would cause very large estimated Doppler frequency differences between the two sites that would be attributed to changes in orbital motion rather than changes in the spacecraft clock. This situation would cause failure of orbit determination and must be avoided.

On the other hand, if the changes in frequency originated on the ground, this points out

an important design consideration of any future low-cost ground station. That is, the GPS synchronization of the ground stations can be done incorrectly for the purposes at hand. The standard method of disciplining an oscillator using a 1 PPS GPS signal is to wrap a feedback control loop around a voltage-controlled oscillator. If this feedback loop is digital, then the system will experience some step changes in frequency. It would be far better to allow the oscillator to free run during data collection and use simultaneously-collected GPS samples to perform the time alignment of the samples in post-processing. Allowing the oscillator to free run prevents the introduction of undesirable step changes in frequency.

In fact, this exact problem was encountered using downlink carrier phase to perform orbit determination for the Iridium satellite constellation. That orbit determination system also relied on multiple ground stations simultaneously observing a satellite in order to perform orbit determination. Initial work in simulation was successful [12], but when on-orbit data were used the orbit determination filter diverged wildly in some situations. It was later discovered that the ground stations had very poorly-disciplined clocks [13], and the orbit determination filter needed to explicitly estimate the behavior of the receiver clocks. When two different clocks had large, unmodeled errors, the filter attributed the difference to orbital motion of the satellite, causing divergence. A similar situation can be avoided in this case by explicitly addressing problems of spacecraft and receiver clock errors prior to deployment of the orbit determination system.

The frequency step magnitudes can be estimated and removed from the data, proceeding under the assumption that the frequency steps originate in the ground equipment. This provides a somewhat more realistic signal detection environment. Estimating and removing the frequency steps is a matter of simultaneously estimating coefficients of a polynomial fit through the frequency history and a series of bias-like step sizes. That is, approximate the frequency history by

$$f_D(t) \approx \sum_{i=0}^{n} c_i t^i + \sum_{j=1}^{m} H(t - t_j) b_j \tag{3.17}$$

where the c_i are the coefficients of a polynomial of order n, $H(t-t_j)$ is the Heaviside function with input argument tailored to yield the 0-to-1 transition at time $t = t_j$, and the m biases b_j are the unknown magnitudes of steps that apply at the known times t_j. That is, the Doppler frequency at time t is the value of a polynomial plus all the biases that apply prior to time t.

When estimating the c_i and b_j, an appropriate order of polynomial for the data set must be chosen. Too low an order of polynomial results in under-fitting the data (actual variation in carrier Doppler frequency will not be well modeled), while too high an order polynomial results in over-fitting the data (the algorithm will fit the noise, and it may model the steps partly using the b_j terms and partly using extra wiggles of the high-order polynomial). In general, shorter spans with fewer steps are better fit. An example of a reasonably good fit is shown in Figure 3.12 using a fifth order polynomial and 7 b_j values to model the 7 steps shown in the figure.

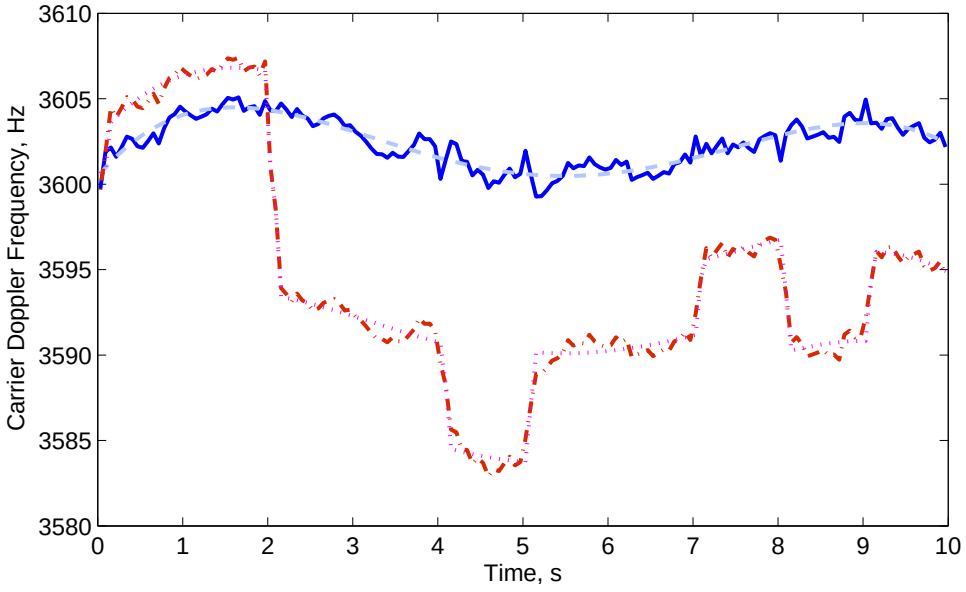


Figure 3.12 – The original carrier Doppler frequency (red dash-dotted) is shown along with the optimized Eq. (3.17) approximation (magenta dotted). The result of subtracting out the estimated biases from the raw data is shown in solid blue along with the polynomial fit with the biases subtracted out.

It is important to note that the result of estimating and removing these frequency biases is not the true frequency history as transmitted by the satellite. The result is merely somewhat more akin to what one would expect to see if neither the spacecraft nor the ground equipment underwent step changes in clock frequency.

Since both the high-SNR and low-SNR channels were sampled and mixed by equipment phase locked to the same 10-MHz reference, both channels should have identical mixing frequencies and sample rates. There is certainly a small amount of inter-channel delay, but this should be small compared to the sample interval, and can be ignored. Because of these facts, easily-acquired knowledge of the signal characteristics derived from the high-SNR channel can be transferred with very high confidence to the low-SNR channel. Of course, an operational system would not have this luxury. It is also hoped, however, that this would not be necessary because an operational system would be designed not to undergo these step changes in frequency.

For the acquisitions presented in the next section, the step changes in frequency were removed. This was accomplished by rotating the samples subsequent to each step change in frequency by the opposite of the integral of the estimated amplitude of the step. The high-SNR channel was used to calculate the optimal signal parameters $\hat{S}_k$, and these were used as a starting point for finding and optimizing the signal parameters for the low-SNR channel.

3.6.2 Results of Acquisition Using the SGH

Three data sets captured during the proof-of-concept experiment used the standard gain horn (SGH) antenna. The estimated C/N_0 from the link budget is about 32.0 dB-Hz. The C/N_0 as measured by the signal analyzer and accounting for modulation suppression of the carrier was between about 30 and 32.5 dB-Hz, depending on the data set. Signal acquisition

in this regime is not trivial, but is not very challenging. Results for this regime are included here because they are very clear and are of high quality. The goal, after all, is to provide measurements to an orbit determination algorithm. The quality of these measurements is discussed in more detail in Chapter 4.

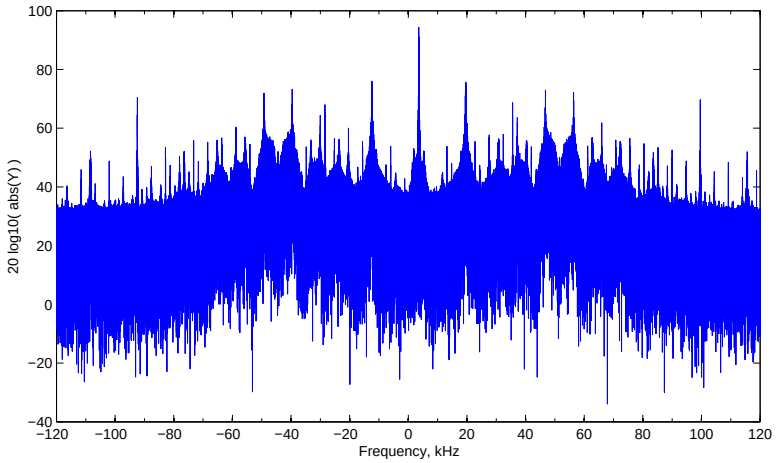
A signal is present in the samples from the SGH, a conclusions that follows from the fact that the spectrum contains a clear peak at the same frequency as observed in the spectrum of the samples from the high-gain antenna over the entire duration of all data batches. Figures 3.9 and 3.10 and also see Figure 3.13 show evidence.

However, it remains to be proven that the signal producing that peak is actually the signal sought. It is possible that other signals are present in the data. The spectrum allocation for spacecraft telemetry signals includes a great deal of frequency reuse [14, 15]. This reuse is possible because highly directional antennas are typically used to receive the telemetry signals. Using a low-gain antenna such as an omnidirectional antenna or even an SGH, significantly reduces directional selectivity, making it probable that other signals with similar carrier frequencies are present in the digital samples.

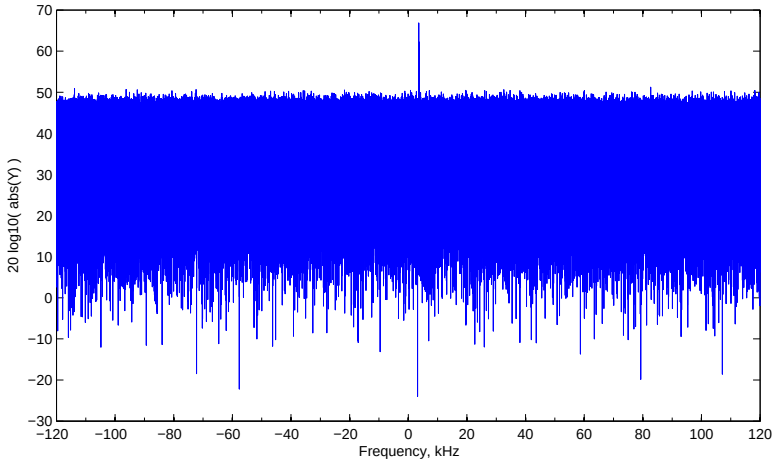
The carrier frequency of the other signals is important. Despite the fact that they would presumably carry very different data, the class of telemetry signal under investigation has a peculiar feature. The phase modulation of the carrier is small enough that it is possible to recover the majority of the carrier power without matching the subcarrier at all. As a consequence, it would be possible to latch onto entirely the wrong signal, simply by matching its carrier frequency.

The remainder of this section seeks to do three things:

- to demonstrate that a signal was detected and quantify the level of certainty that a signal is present
- to demonstrate that the signal acquired was the signal that was sought



(a) High-SNR from 13-m parabolic antenna.

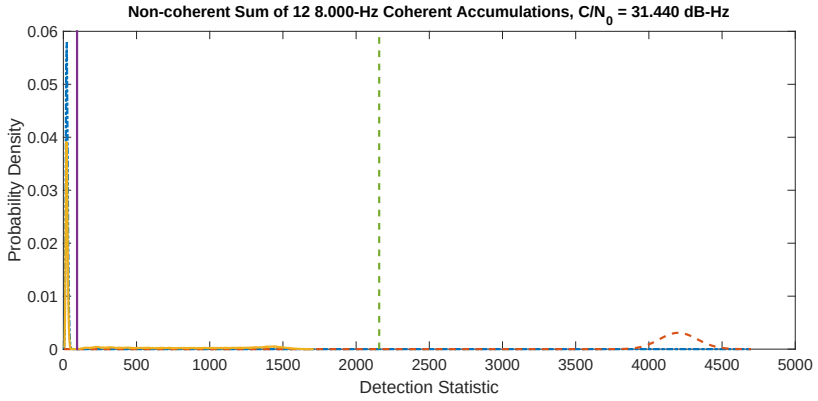


(b) Low-SNR from SGH antenna.

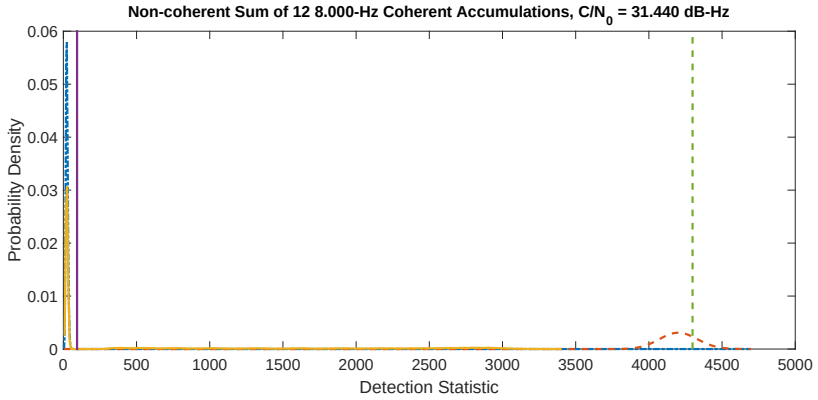
Figure 3.13 – The two figures show the discrete Fourier transform (DFT) power spectra of samples recorded from the high-and low-SNR signal paths. A clear carrier peak is visible in the low-SNR spectrum at the same 3.5 kHz Doppler shift as in the high-SNR spectrum.

- and to show why it is possible to latch onto the wrong signal.

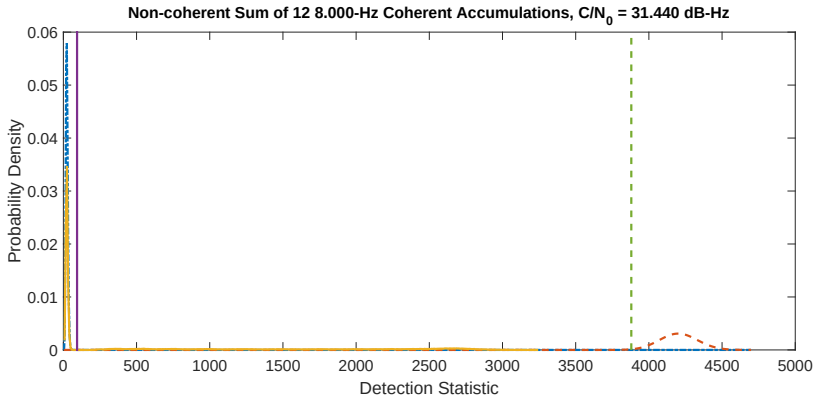
For the first topic—demonstrating signal detection—the hypothesis test described in Section 3.3 is the primary tool. Figure 3.14 shows one detection test for each of the three SGH data sets. Each of the figures displays two probability density functions (PDFs) as well as a histogram of accumulations calculated using the on-orbit data. The PDF on the left shows the theoretical probability density of accumulations computed on noise only, referred



(a) Data set 1.



(b) Data set 2.



(c) Data set 3.

Figure 3.14 – The figures show the noise-only PDF (blue dashed, at left), the signal-and-noise PDF (red dashed, at right), the histogram of actual computed accumulations (solid yellow, and likely to be mostly signal-not-present accumulations due to incorrect guess of $\hat{S}_k$ for most of them), the detection threshold (vertical solid purple), and the highest accumulations after the grid search (vertical dashed green). In each detection, $T_{coh} = 0.125$ s and $M = 12$ coherent accumulations per non-coherent accumulation. All three data sets were recorded with the SGH antenna.

to as $p(x|H_0)$ in Section 3.3. Accumulations that are computed with signal parameters that are statistically independent of the true parameters do appear to come from a distribution that closely matches $p(x|H_0)$. The histograms, shown in solid yellow, all have a large peak that falls under the theoretical $p(x|H_0)$. They fall *under* $p(x|H_0)$ because the histograms are normalized to the total area under the curve, and a large number of accumulations are drawn from distributions that are not $p(x|H_0)$. This is discussed more later in this section, and the next section will feature a similar figure, but with histograms that are easier to see.

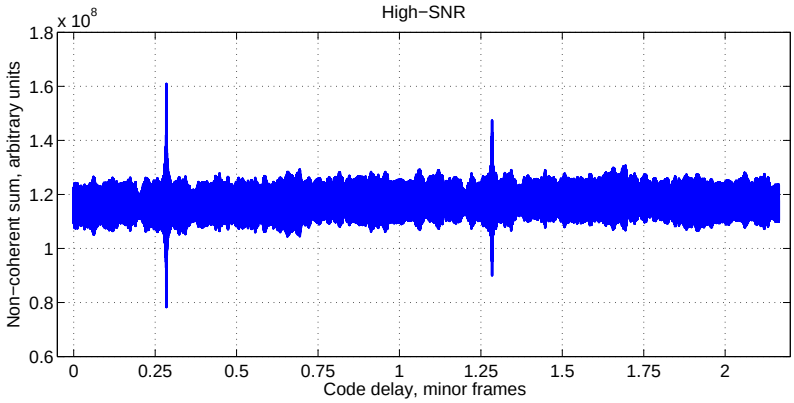
The PDF on the right is the theoretical probability density of accumulations of signal and noise. That is, if the signal replica matched the signal exactly, except for the presence of additive, white, Gaussian measurement noise, the resulting accumulation would be drawn from the PDF on the right.

In practice, the model does not match the actual signal perfectly due to unmodeled carrier and subcarrier phase noise processes, and accumulations tend to be smaller than predicted. The vertical lines in Figure 3.14a–c show the detection threshold (solid purple) and the highest accumulation from the grid search (dashed green). The detection threshold γ_F is set such that probability of false alarm α_F is 1.0×10^{-5} according to Section 3.3 using $N_S = 1.38 \times 10^6$ statistically independent detection test statistics. The highest accumulations in each subfigure of Figure 3.14 are far to the right of the detection threshold. This indicates three positive detection tests. That is, in each case, we can declare with confidence that we have found a signal. In fact, if the detection threshold were set to the accumulation value shown in Figure 3.14(a), the probability of any one sample from $p(x|H_0)$ reaching or exceeding that level is about 8×10^{-443}, which is effectively statistically certain never to occur. In other words, in order to attain the accumulations shown in Figure 3.14, there *must* be signal present. The relatively low maximum accumulation of Fig. 3.14(a) with respect to the signal-present PDF may be due to a poor choice of $\hat{S}_k$, especially τ_{B0} may be incorrect by an integer multiple of minor frames, or what is more likely is that the signal-present PDF is

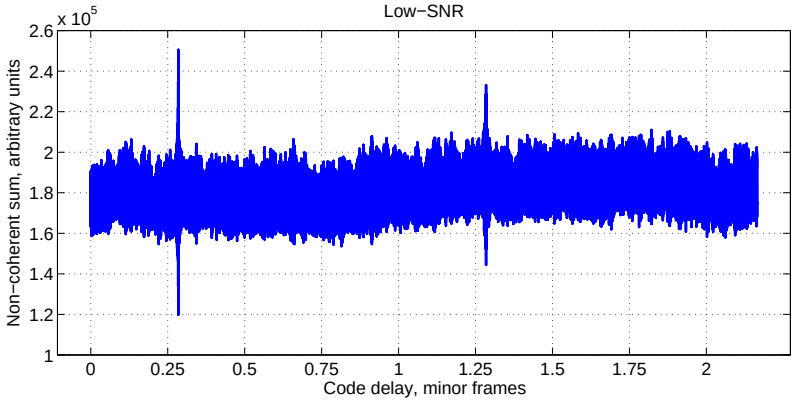
too high. The mean of the signal-present PDF is calculated *a priori* from the C/N_0 expected from the link budget.

It may be certain that *a* signal has been found, but it still must be demonstrated that the signal found was the signal sought. There are two signs that will make this certain. If the signal that the detection algorithm found was indeed the signal that it was looking for, then correctly mixing off the subcarrier should have produced a higher accumulation than incorrectly mixing off the subcarrier. Moreover, the minor frames transmitted by the satellite are largely the same from one minor frame to the next (more than 50% of the bits are identical), at least over a span of a few minor frames, so there should be a strong periodicity in the correlation of the signal replica with the true signal. Figure 3.15 demonstrates both of these signs for both the high-SNR and low-SNR signal paths. Not only does the accumulation value increase when the subcarrier is mixed correctly (*i.e.*, at the right code offset), but the expected periodicity (at a separation of exactly one telemetry frame) appears. This indicates strongly that the correct signal has been found.

A skeptic might say that these periodic increases could result from correlating the signal with any periodic string of bits. In order to satisfy this skeptic, correlations of the signal with other, incorrect bit streams have been calculated. Figure 3.16 shows some typical results for the low-SNR samples taken from the SGH. Figure 3.16(a) shows correlations of the samples from the SGH with a replica signal generated using the data stream from the Cortex modem—the correct data stream—as the code offset is varied. All other parameters are the optimal parameters. The mean value of the wide, horizontal band is effectively due to correlation of the carrier only, and its width is proportional to how many of the bits in the data stream correlate. The high, narrow peaks occur when the code replica is closely aligned with the actual received code, plus or minus an integer number of minor frames. The strong anti-correlations occur half of a subcarrier cycle on either side of the peaks, so that most of the subcarrier is anti-correlated. A detail of this behavior is shown in Figure 3.17.



(a) High SNR.



(b) Low SNR.

Figure 3.15 – Accumulations on the vertical axis are shown versus code delay on the horizontal axis for both the high- and low-SNR paths. The low-SNR path used the SGH antenna. A strong periodicity with period of exactly one minor frame is visible. There is also a periodicity due to individual subcarrier cycles which is too fine to be seen on this scale.

The results of computing the same correlations but against three different incorrect data streams are shown in Figures 3.16(b) to 3.16(d). These figures feature the wide band due to the carrier power, as in Figure 3.16(a), but they conspicuously lack the strong peaks at minor frame intervals. The varying widths of the bands are due to differing numbers of bits matching for the different bit streams. It is interesting to note that the actual bit stream contained 199.5% as many '0' bits as '1' bits (121403 '0's v. 60869 '1's in the entire 38-second data batch) and the variation in accumulation was correlated with the number of bits that matched. The variation is greater for the all-zeros stream than it is for the all-ones stream.

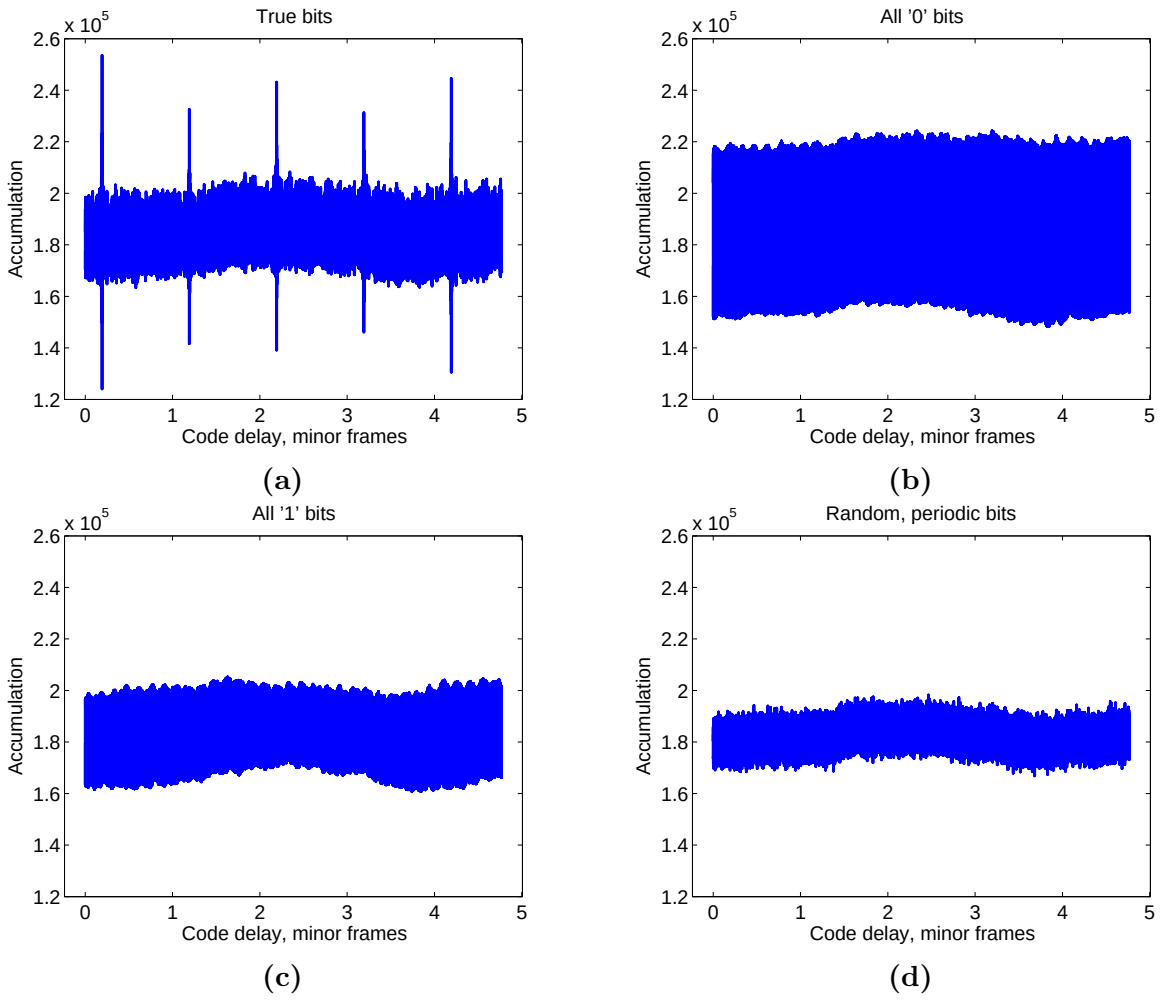


Figure 3.16 – Correlation versus code delay in minor frame periods for samples from the SGH. All of these correlations were calculated using the correct carrier and subcarrier parameters but the data streams differed. In (a), the correct data stream was used. In (b), the data stream contained all '0' bits. In (c), the data steam contained all '1' bits. And in (d), the data stream was a repeated string of 2048 random bits.

Correlations were also computed for many other incorrect bit streams in addition to those shown in the figures. Some of the bit streams were periodic with period equal to one minor frame, and some had no periodicity imposed. They all resulted in correlations similar to those shown in Figure 3.16(d). After seeing the results shown in these figures, it is hoped that the hypothetical skeptic would be convinced that the true signal was in fact found in the samples from the SGH.

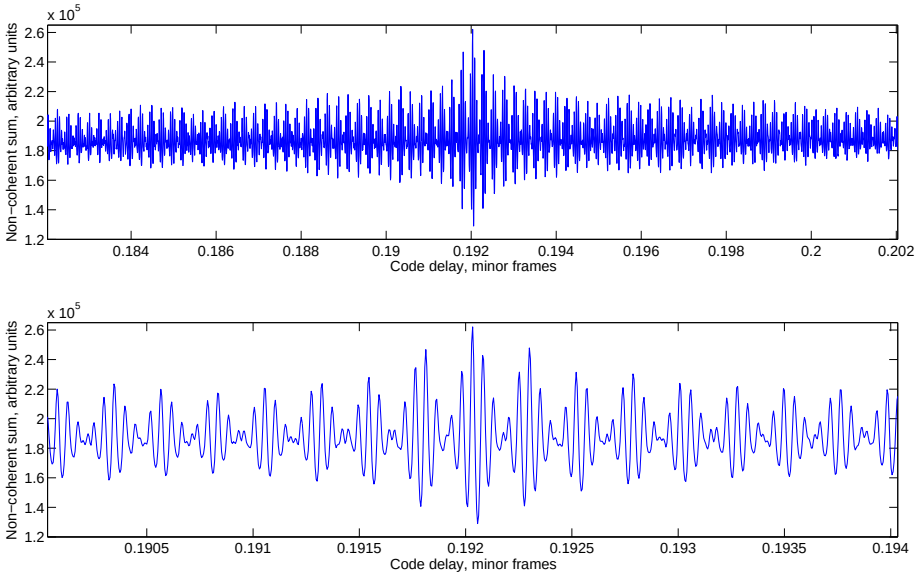


Figure 3.17 – These are two detail views of Figure 3.16(a) centered on the first large correlation peak. The high-frequency variation visible in the lower panel occurs at the subcarrier frequency. That is, there is a periodicity due to individual subcarrier cycles. The longer-period variations have a period equal to the Manchester chip width, *i.e.*, 5 subcarrier cycles.

The last topic of this section is to show why, if other signals are present, it would be possible to find the wrong signal. Some of the answer to this question has already been discussed tangentially. When discussing Figure 3.14, it was mentioned that there are many accumulations that are drawn from a distribution that is neither $p(x|H_0)$ nor $p(x|H_1)$. It was also mentioned when discussing Figure 3.16 that the mean value of the accumulations is independent of the bit stream the signal is mixed against. By referring to one more figure, these two facts can be brought together. Figure 3.18 shows accumulation value on the vertical axis (and in color) versus a two-dimensional cross section of the signal parameter space. It shows code offset on one horizontal axis and carrier Doppler frequency on the other. The other signal parameters are held constant. Accumulations drawn from the $p(x|H_0)$ distribution are the dark blue floor of the figure. The one accumulation drawn from $p(x|H_1)$ is the highest point on the peak near the left side of the figure. All the other accumulations, those that make up the ridge in the figure, are drawn from neither distribution. These are

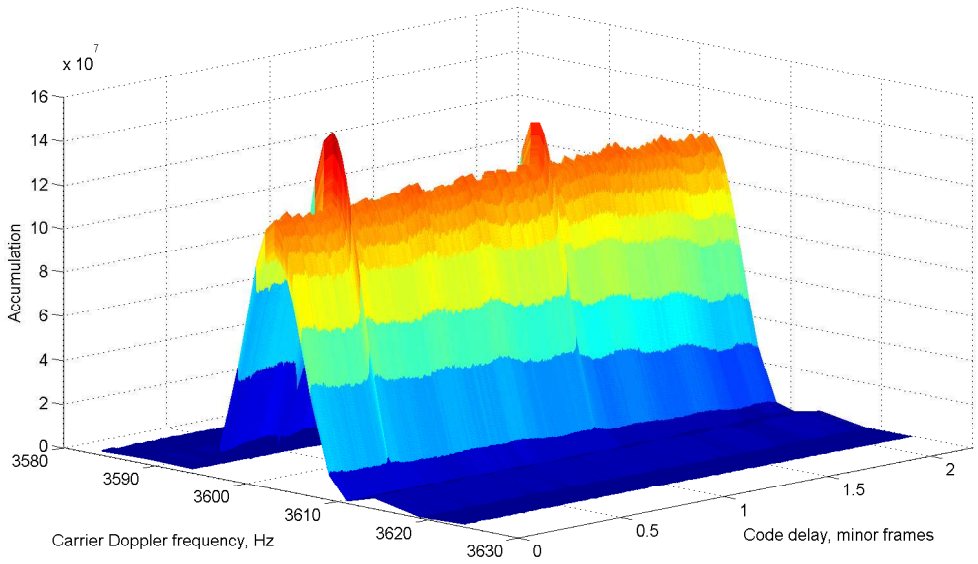


Figure 3.18 – The figure shows non-coherent accumulation value on the vertical axis versus both carrier Doppler shift and code delay on the two horizontal axes.

a by-product of the phase modulation of the telemetry signal. Referring to Figure 3.1, it is clear that simply correctly mixing off the carrier produces an accumulation that is much closer to the value that would be attained by also correctly mixing the subcarrier than it is to zero. This corresponds to the average value of the top of the ridge in Figure 3.18. In fact, Figures 3.15 and 3.16(a) are essentially slices along the peak of the ridge in Figure 3.18, but for a different data set. One can imagine that a figure similar to Figure 3.18 could be constructed for correlations against an arbitrary and incorrect bit stream, such as the one used to construct Figure 3.16(d). The result would be a ridge with the same height as in Figure 3.18 but without the fins that result from correlating against the correct bit stream at the correct code delay. Achieving an accumulation higher than the floor of the figure indicates that *a* signal is present. The presence of the fins at minor frame intervals indicates that *the* signal is present.

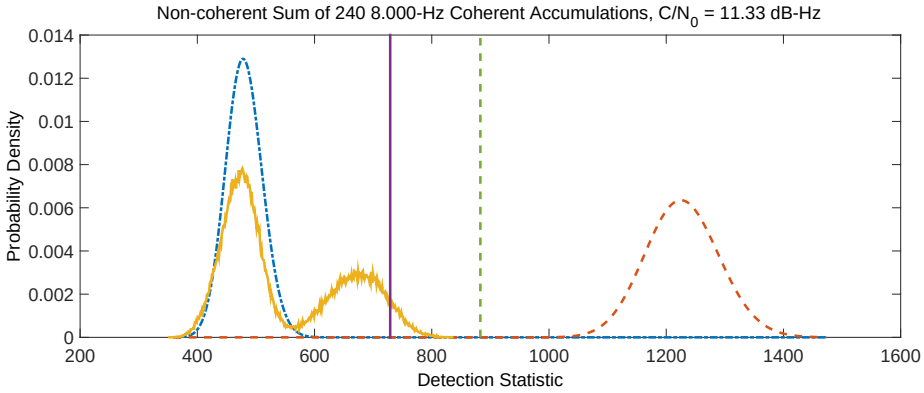
The results in this section all use 12 coherent sums each computed over one eighth of a second, or 1.5 seconds of samples total. The signal can be detected using many more

combinations of number and duration of coherent sums, and several have been tried with positive results. It is also possible to reliably detect the signal with less than 1.5 s of data at the C/N_0 observed, but as the coherent interval decreases and the number of coherent intervals decreases, the corresponding estimates of the carrier Doppler observable are of poorer quality. The quality of the Doppler measurements will be discussed in the next chapter. The next section presents results of the acquisition calculations using data from an omnidirectional antenna.

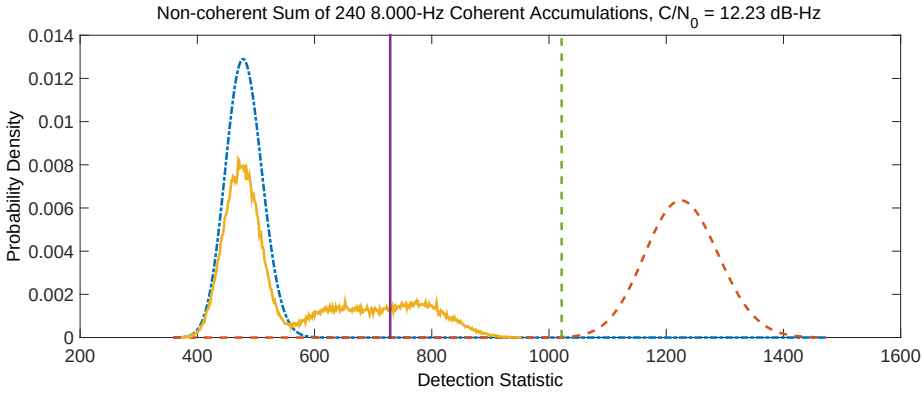
3.6.3 Results of Acquisition Using the Omni

Three data sets captured during the proof-of-concept experiment used the omnidirectional antenna. The estimated C/N_0 for these data sets was 11.0 dB-Hz from the link budget in Table 2.1. No estimate could be made from the signal analyzer, as the entire signal was buried under the noise floor. The signal from the omnidirectional antenna was significantly weaker than from the SGH. Acquisition at this level of C/N_0 is quite challenging.

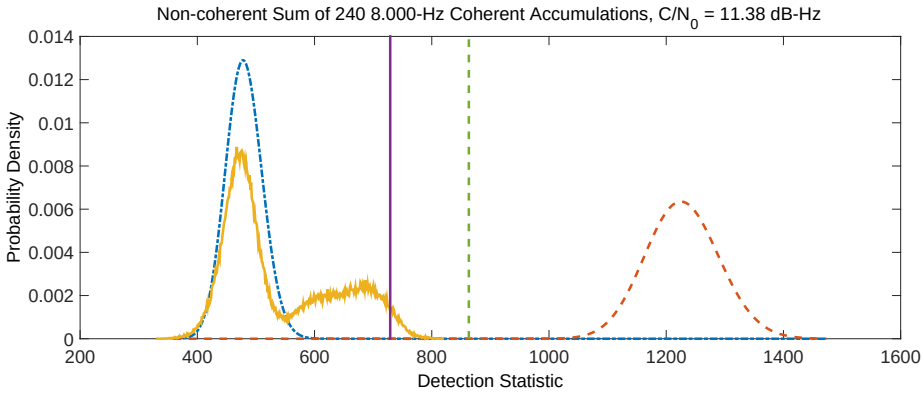
Nonetheless, the signal can be acquired in the digital samples. After acquisition, the C/N_0 was determined based on the estimated signal amplitude to be between 11 and 12 dB-Hz, depending on the data set. Note that in the detection tests shown below, the estimated abrupt frequency increments shown in Figs. 3.7 and 3.8 of Section 3.6.1 were mixed out prior to acquisition. Figure 3.19 depicts one detection test for each of the three data sets captured using the omnidirectional antenna. The three omnidirectional antenna detection tests each used 30 seconds of data in 240 coherent integrations each of one-eighth of a second duration. In each of these three detection tests, the detection threshold was exceeded, but not by nearly as large a margin as with data from the SGH. These are positive detections, but they are not nearly as strong. Notice that the range of the horizontal axis of Figure 3.19 is small enough that the forms of the two PDFs and the histogram of the calculated accumulations



(a) Data set 4.



(b) Data set 5.



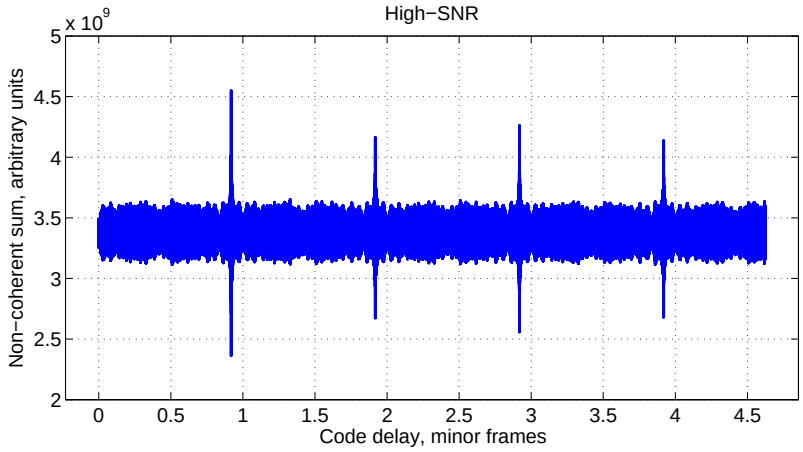
(c) Data set 6.

Figure 3.19 – The figures show the noise-only PDF (blue dashed, at left), the signal-and-noise PDF (red dashed, at right) for a signal with C/N_0 as estimated from the link budget, the histogram of accumulations (solid yellow), the detection threshold (vertical solid purple), and the highest accumulations after the grid search (vertical dashed green). In each detection, $T_{coh} = 0.125$ s and $M = 240$ coherent accumulations per non-coherent accumulation. All three data sets were recorded with the omnidirectional antenna. The histogram of computed accumulations looks so odd because it is being strongly affected by the presence of the signal. Accumulations with incorrect τ_{B0} straddle the detection threshold.

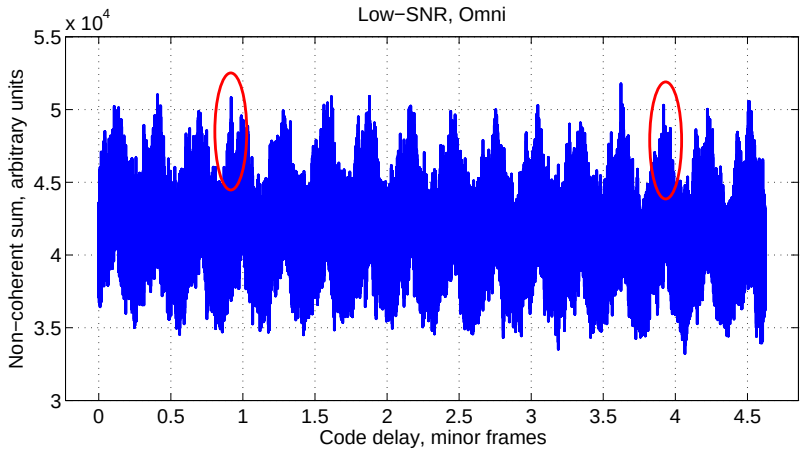
are clearly discernible, as opposed to those from the SGH detections in Figure 3.14.

At these levels of C/N_0, the increase in accumulation due to correct alignment of the subcarrier is less than the variation in accumulation due to measurement noise. As a result, the code start time measurement is not reliable at this C/N_0. This is illustrated in Figure 3.20. In the figure, the red ellipses in the lower panel highlight correlation peaks that are probably due to alignment of the subcarrier. However, there is another peak, likely a noise-induced peak, that is larger than either of these two peaks. Other data sets collected using the omni-directional antenna exhibit similar behavior—that is, either no discernable peaks due to subcarrier alignment, or peaks that are smaller than variation due to measurement noise.

In the accumulations computed from data captured using the omnidirectional antenna and shown in Figure 3.19, there is very little evidence that the signal found was actually the signal sought. In fact, the only evidence is that the carrier Doppler shift agrees well with the carrier Doppler shift measured from the 13-m antenna. This does give an indication of success, but it is not an unambiguous indication. It is possible that another satellite's signal, which arrived at the antenna with the same observed carrier frequency, could be mistaken for the signal from the test satellite. This could happen because the subcarrier contributes so little to the overall power of the detection. In order to underscore this fact, a detection test was run using entirely the wrong string of bits. The bits used were a repeated string of 2048 random bits (the length of one telemetry frame). The signal was detected, albeit at a slightly lower level than when using the correct code. The highest accumulation found during the grid search using the incorrect code was 802.30. For the same data set using the correct code, the highest accumulation was 813.57 (prior to the optimization step described in Chapter 4)—this detection is shown in Figure 3.19(a) as the right-most end of the blue curve. The detection threshold for this data set was 729.09. Using the omnidirectional antenna, the signal was detected, but at such a low SNR that it was impossible to distinguish the



(a) High SNR.



(b) Low SNR.

Figure 3.20 – Accumulations on the vertical axis are shown versus code delay on the horizontal axis for both the high- and low-SNR paths. The low-SNR path used the omnidirectional antenna. The high-SNR accumulations clearly show the correct alignment of the subcarrier, while the low-SNR accumulations do not. Two red ellipses are superimposed on the figure where peaks occur at the correct code delay.

subcarrier and its data content. Effectively, only the carrier was being detected.

Possible methods of detecting subcarrier alignment include increasing the coherent accumulation interval or increasing the non-coherent interval. Unfortunately, both of these methods involve using a longer time span of data. Since, in the current model, the subcarrier is modeled as continuous and coherent with only a quadratic phase variation (linear frequency variation), longer accumulations are more affected by phase noise of the subcarrier-generating spacecraft clock. That is, over too long an interval, small variations in the subcarrier phase

tend to wash out the accumulated power rather than add to it. This appears to be the case for the collected data. Using higher-quality frequency sources might mitigate this effect, allowing for longer coherent integrations.

Despite the fact that the subcarrier was not detected using the omni-directional antenna, the measured carrier Doppler shift was of reasonable quality. This may or may not be acceptable during operation. A tradeoff between measurement quality and system complexity is clearly evident. A standard gain horn antenna, which must be pointed to within a few degrees of the target bearing produces much better measurements than an omni-directional antenna, which needn't be pointed in any special direction.

3.6.4 Computational Burden

After having addressed the question of whether or not it is possible to find the signal, it is now necessary to consider the computational practicality of finding the signal. In the above analysis, the search for the low-SNR signal was always heavily aided by prior results from simultaneously recorded high-SNR samples. Table 3.1 displays information relevant to the computational time required by the signal acquisition calculations for the low-SNR signal with reduced search space due to aiding in carrier and subcarrier Doppler from the high-gain antenna. The times listed in the table were calculated running the signal acquisition software in MATLAB under Windows 7 on all four cores of an Intel Core i7 processor running at 2.9 GHz. The times listed reflect the time necessary to process the low-SNR samples with aiding from the high-SNR signal. The aiding provided collapsed the carrier and subcarrier Doppler dimensions. However, due to the recording start time uncertainty on the two signal chains, there was no aiding in the code delay dimension provided by the high-SNR signal to the low-SNR signal.

In order to perform a signal acquisition without aiding, it would be necessary to search

Table 3.1 – Summary of computational burden of signal search.

SGH	
Coherent accumulation length, samples	62500
Number of coherent accumulations	4
Number of code offsets	262144
Time to compute coarse acquisition, s	0.313±0.035
Time to compute fine acquisition, s	4.344±0.077
Omni	
Coherent accumulation length, samples	500000
Number of coherent accumulations	50
Number of code offsets	1048576
Time to compute coarse acquisition, s	22.6±2.1
Time to compute fine acquisition, s	453.8±1.3

the entirety of both Doppler dimensions and to search over the entire range of possible code delays. The range of carrier Doppler shifts that might be encountered is affected by both motion-induced Doppler shift and carrier oscillator frequency error. While the Doppler shift due to relative motion should be well-known at the time of signal acquisition, the carrier clock's frequency error may not be known. A discussion of Doppler grid spacing may be found in Section 3.4. In short, the carrier clock is specified to have maximum frequency error of no more than 1.0 part-per-million. Operating in K_u band at 12.6985 GHz, as in the proof-of-concept experiment, that equates to ±12.6985 kHz of carrier frequency uncertainty. This entire range must be searched during an unaided signal acquisition. Assuming coherent accumulations of 1.0 s duration, as for the detections calculated using the omni antenna, the carrier Doppler grid spacing must be no more than 0.5 Hz, resulting in 50794 carrier Doppler grid points.

Likewise, the subcarrier-generating clock is specified to have frequency error of no more than 50 parts-per-million. The subcarrier frequency is 48 kHz. The range of possible subcarrier frequency errors is only ±2.4 Hz. However, since the subcarrier is modeled as coherent over the entire detection interval, the tolerable frequency error is very small. Assuming that a detection interval is 50 s long, as used for detecting the signal from the omni antenna, the

subcarrier Doppler grid spacing must be no more than 0.01 Hz, resulting in 480 subcarrier Doppler grid points.

If the time uncertainty were the same as assumed during the detections calculated using data from the omni, 2.1 s, then the total time to compute an acquisition would be the time in Table 3.1, 22.6 s, multiplied by the product of the numbers of carrier and subcarrier Doppler grid points, or a total of a staggering 17.46 years. Under the signal detection parameters used for the SGH data, the search time is reduced to a much more reasonable 2.8 hours. But even 2.8 hours is not practical.

There are many ways to speed the grid search. Perhaps the most practical way would be to aid the initial signal acquisition with signal parameters measured using a high-gain antenna and an estimate of the current orbital parameters. This would allow the system to determine the spacecraft clock frequency error and the relative-motion-induced Doppler shift. This information could be propagated to the other ground stations to reduce their search space. Such a system would result in acquisition times of only a few times those shown in the table.

Another method of making this search practical would be to modify future spacecraft hardware such that the subcarrier clock is slaved to the carrier clock, or such that the subcarrier and carrier are generated phase coherently from a single frequency source. This would eliminate the subcarrier Doppler search dimension entirely. Failing this, it would be possible to include a small electronic device on the spacecraft that measured the frequency drift between the carrier and subcarrier and reported this number as part of the telemetry message. Once initial acquisition was complete, probably using a higher-gain antenna, the relative drift information could be passed along to the ground stations to reduce their subcarrier Doppler search spaces.

Timing uncertainty must also be considered. During the proof of concept experiment, the

timing uncertainty was due to the difference in the start times of the recordings of the two data streams—there was no way to put absolute time tags on the samples. In an operational system, the timing uncertainty would be due to uncertainty in the position of the satellite and due to spacecraft and receiver clock error. If orbit information were provided to all of the ground stations in the system, they could use this information to reduce the timing uncertainty due to spacecraft position uncertainty. If a pseudo random number (PRN) sequence were transmitted in place of a telemetry message, then the timing uncertainty due to spacecraft clock error could be limited to modulo the duration of the PRN sequence. This alone could provide an enormous signal search time saving. The receiver clocks could be aided by a high-quality frequency source, like GPS, to provide absolute time tags for the samples. All these things in combination would reduce the timing uncertainty to a matter of milliseconds.

Table 3.2 shows the projected search times for reasonable levels of residual uncertainty for an aided signal detection. The projected search time when searching for the signal in samples from a SGH is entirely acceptable, a total of just over 7 s. The total time using an omni antenna is still too high to be practical without better aiding.

Speed increase could also be realized by a switch from MATLAB to a lower-level language, like C or C++. Some anecdotes suggest that speed increases of 10–100 times can be expected[2]. This alone would not yield enough improvement to make acquisitions using an omni antenna practical, except in heavily-aided scenarios, but it might allow unaided or minimally-aided acquisitions using a slightly higher-gain antenna, such as an SGH or a small parabolic reflector like those used for satellite TV reception.

It is computationally practical to acquire the signal from an SGH. Though it is possible to acquire the signal from an omni antenna, it is not practical without significant aiding.

[2] *E.g.* this thread http://stackoverflow.com/a/24643248 and the linked answer specifically suggest that an average speed increase of 50 times when moving from MATLAB to C++ is reasonable.

Table 3.2 – Projected search time for aided acquisitions.

Assumed Uncertainty	
Spacecraft position uncertainty, m	2000
Spacecraft velocity uncertainty, cm/s	30
Carrier clock frequency error uncertainty, Hz	2
Subcarrier clock frequency error uncertainty, Hz	0.5
Subcarrier clock error uncertainty, s	0.01
Receiver clock error uncertainty, s	1×10^{-9}
SGH	
Number of carrier Doppler grid points	7
Number of subcarrier Doppler grid points	3
Number of code offsets	16384
Coherent accumulation length, samples	62500
Number of coherent accumulations	4
Time to compute coarse acquisition, s	2.823
Time to compute fine acquisition, s	4.237
Omni	
Number of carrier Doppler grid points	59
Number of subcarrier Doppler grid points	100
Number of code offsets	8192
Coherent accumulation length, samples	500000
Number of coherent accumulations	50
Time to compute coarse acquisition, s	71330
Time to compute fine acquisition, s	466.5

The next chapter discusses fine acquisition and the quality of the Doppler measurements that can be obtained from acquisitions calculated from both an SGH and an omni antenna.

CHAPTER 4

FINE ACQUISITION AND EXTRACTION OF OBSERVABLES

Coarse acquisition provides only a rough idea of the signal parameters—they are accurate to within one half the grid spacing used in the grid search. In order to find more accurate signal parameters, the parameter vector which results from a coarse acquisition can be used as the starting point for a non-linear optimization that finds the signal parameters which minimize the squared-error in the signal model as in Eq. (3.4), or equivalently, which maximizes the detection statistic (the properly normalized non-coherent accumulation) in either Eq. (3.3) or (3.12). Any suitable non-linear optimization technique could be used. The technique described below is a modified Gauss-Newton optimization[16]. The method is different from a standard Gauss-Newton optimization in that only a subset of the parameters are explicitly optimized.

Following the Gauss-Newton optimization, the fine acquisition is complete, and signal observables may be calculated. In this case, we will consider the Doppler shift observable in detail, after fine acquisition is dealt with.

4.1 Gauss-Newton Optimization

The fine acquisition algorithm consists of a Gauss-Newton optimization routine which minimizes an objective function with respect to the signal parameter vector. The following algorithm description is largely a standard Gauss-Newton algorithm, as described in Reference [16], but with notation adapted to the current problem. In this case, the objective

function, f, is half the sum of the squares of the residuals for the entire non-coherent interval.

$$\begin{aligned} f(\boldsymbol{x}) &= \frac{1}{2}\sum_{m=1}^{M} \left(\boldsymbol{y}_m - \hat{\boldsymbol{y}}_m\right)^{\dagger} \left(\boldsymbol{y}_m - \hat{\boldsymbol{y}}_m\right) \\ &= \frac{1}{2}\sum_{m=1}^{M} \bar{\boldsymbol{r}}_m^{\dagger} \bar{\boldsymbol{r}}_m \\ &= \frac{1}{2}\sum_{m=1}^{M} \left(\boldsymbol{r}_{Im}^{\mathrm{T}} \boldsymbol{r}_{Im} + \boldsymbol{r}_{Qm}^{\mathrm{T}} \boldsymbol{r}_{Qm}\right) \end{aligned} \tag{4.1}$$

Above, the parameters to be optimized are $\boldsymbol{x}$=[α_{Ak}, ω_{DAk}, α_{B0}, ω_{DB0}, τ_{B0}, X_1, $Y_1, \ldots, X_M$, Y_M]$^{\mathrm{T}}$, $\boldsymbol{y}_m$ is the vector of complex-valued signal samples to be used in the m^{th} coherent integration, $\hat{\boldsymbol{y}}_m$ is the vector of complex-valued replica signal samples in the m^{th} coherent integration based on the parameters $\boldsymbol{x}$, the dagger superscript (†) is the complex conjugate transpose, $\bar{\boldsymbol{r}}_m = \boldsymbol{y}_m - \hat{\boldsymbol{y}}_m$ is the complex-valued vector of residuals for the m^{th} coherent integration, and its real-valued components in the real and imaginary directions are $\boldsymbol{r}_{Im}$ and $\boldsymbol{r}_{Qm}$, respectively. The vectors of all the real and imaginary residuals for all coherent integration intervals are $\boldsymbol{r}_I = [\boldsymbol{r}_{I1}^{\mathrm{T}}, \boldsymbol{r}_{I2}^{\mathrm{T}}, \ldots, \boldsymbol{r}_{Im}^{\mathrm{T}}, \ldots, \boldsymbol{r}_{IM}^{\mathrm{T}}]^{\mathrm{T}}$ and $\boldsymbol{r}_Q = [\boldsymbol{r}_{Q1}^{\mathrm{T}}, \boldsymbol{r}_{Q2}^{\mathrm{T}}, \ldots, \boldsymbol{r}_{Qm}^{\mathrm{T}}, \ldots, \boldsymbol{r}_{QM}^{\mathrm{T}}]^{\mathrm{T}}$, and these two vectors can be stacked into one $\boldsymbol{r} = [\boldsymbol{r}_I^{\mathrm{T}}, \boldsymbol{r}_Q^{\mathrm{T}}]^{\mathrm{T}}$. That is, the unsubscripted $\boldsymbol{r}$ is the real-valued vector of residuals in the real direction for all M intervals stacked on the real-valued vector of residuals in the imaginary direction for all M intervals.

The Gauss-Newton method relies on having approximations of the first and second derivatives of f with respect to the parameter vector components $[\boldsymbol{x}]_i = x_i$ [16]. The derivatives can be expressed in terms of the $2p \times n$ Jacobian matrix of the residuals with respect to the parameters, $J(\boldsymbol{x})$. If J_I and J_Q are the matrices whose (i, j) elements are

$$\left[J_I(\boldsymbol{x})\right]_{i,j} = \frac{\partial r_{I,j}}{\partial x_i} \tag{4.2}$$

and

$$\left[J_Q(\boldsymbol{x})\right]_{i,j} = \frac{\partial r_{Q,j}}{\partial x_i} \tag{4.3}$$

respectively, where $r_{I,j}$ is the j^{th} element of $\boldsymbol{r}_I$ and $r_{Q,j}$ is the j^{th} element of $\boldsymbol{r}_Q$, then

$$J(\boldsymbol{x}) = \begin{bmatrix} J_I(\boldsymbol{x}) \\ J_Q(\boldsymbol{x}) \end{bmatrix} \tag{4.4}$$

The required derivatives can be written

$$\nabla f(\boldsymbol{x}) = J(\boldsymbol{x})^{\mathrm{T}} \boldsymbol{r} \tag{4.5}$$

and

$$\nabla^2 f(\boldsymbol{x}) \approx J(\boldsymbol{x})^{\mathrm{T}} J(\boldsymbol{x}) \tag{4.6}$$

At this point, a series of linear least squares problems is solved, each one successively moving closer to the optimal solution to the non-linear problem. The linear problems to solve are

$$\min_{\boldsymbol{p}} \| J_k \boldsymbol{p} + f_k \|^2 \tag{4.7}$$

and the solutions are

$$\boldsymbol{p}_k = -(J_k^{\mathrm{T}} J_k)^{-1} J_k^{\mathrm{T}} \boldsymbol{r}_k \tag{4.8}$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \kappa \boldsymbol{p}_k \tag{4.9}$$

for some scalar step length $0 < \kappa \leq 1$ that satisfies the Wolfe condition for sufficient objective function decrease, $f(\boldsymbol{x}_k + \kappa \boldsymbol{p}_k) \leq f(\boldsymbol{x}_k) + c_1 \kappa \nabla f_k^{\mathrm{T}} \boldsymbol{p}_k$. If the problem were linear, κ would always be 1. Nonlinearities may force the algorithm to take shorter steps to ensure that the step results in a sufficient cost decrease. The subscript k refers to the k^{th} iteration of the Gauss-Newton method. Of course, $\boldsymbol{x}_0$ is the parameter vector that results from the coarse acquisition. Eventually, $\boldsymbol{x}_k$ and $\boldsymbol{x}_{k+1}$ become nearly identical, their objective function values are also nearly identical, and the best κ value will be 1 or nearly so. At this point, the solution has converged. During these iterations, the step direction $\boldsymbol{p}$ is calculated using a matrix pseudo-inverse. It is important during matrix inversion to properly normalize the equations

in order to maintain adequate numerical conditioning. In the author's experience with the algorithm, a row normalization of $J_k^{\mathrm{T}} J_k$ is adequate to maintain numerical conditioning.

The matrix J_k may be of very large dimension and impractical to store in computer memory. It has $2p$ rows, where $p = MN$ is the number of complex-valued samples in all M coherent integrations, and $n = 5 + 2M$ columns, where M is the number of coherent integrations summed in the non-coherent integration of Eq. (3.3). M may be on the order of 10 to 100 (the examples in Sections 3.6.2 and 3.6.3 used $M = 12$ and $M = 240$, respectively) and p will be on the order of 10^6 to 10^7. However, J_k need never be calculated in full, instead the much more manageably sized $J_k^{\mathrm{T}} J_k$ and $J_k^{\mathrm{T}} \boldsymbol{r}_k$ may be calculated in parts. These matrices are of dimension only $n \times n$ and $n \times 1$. Further, we need not use the nonlinear optimization to explicitly optimize for the last $2M$ elements of $\boldsymbol{x}$, the elements $X_{k(1)}, Y_{k(1)}, \ldots, X_{k(M)}, Y_{k(M)}$, because these parameters enter the problem linearly, and can be estimated using the linear estimation described in Section 2.2.

Both $J_k^{\mathrm{T}} J_k$ and $J_k^{\mathrm{T}} \boldsymbol{r}_k$ can be written as a sum over all coherent integration intervals

$$J_k^{\mathrm{T}} J_k = \sum_{m=1}^{M} J_{m,k}^{\mathrm{T}} J_{m,k} \tag{4.10}$$

$$J_k^{\mathrm{T}} \boldsymbol{r}_k = \sum_{m=1}^{M} J_{m,k}^{\mathrm{T}} \boldsymbol{r}_{m,k} \tag{4.11}$$

where $J_{m,k}$ is the Jacobian of the residuals in the m^{th} coherent integration interval with respect to the parameter vector of the k^{th} Gauss-Newton iteration, and $\boldsymbol{r}_{m,k}$ is the vector of residuals for the m^{th} coherent integration interval of the k^{th} Gauss-Newton iteration. Calculating these quantities this way reduces the maximum number of rows in the matrices involved from $2p = 2MN$ to $2N$. One could break these calculations down even further so that the sums were over each residual element, or even further so that each element of $J_k^{\mathrm{T}} J_k$ or $J_k^{\mathrm{T}} \boldsymbol{r}_k$ was calculated individually as a sum over the residual vector elements. However, this turns out to be unnecessary.

Since we are interested in only the first five rows and columns of $J_k^{\mathrm{T}} J_k$ and only the first five rows of $J_k^{\mathrm{T}} \boldsymbol{r}_k$—those associated with the signal parameters α_{Ak}, ω_{DAk}, α_{B0}, ω_{DB0}, and t_{B0}—we can solve a smaller optimization. Start by breaking $J_{m,k}$ into four blocks,

$$J_{m,k} = \begin{bmatrix} A_{Im,k} & B_{Im,k} \\ A_{Qm,k} & B_{Qm,k} \end{bmatrix} \tag{4.12}$$

The two 'A' blocks on the left are each $N \times 5$ matrices. These are elements of $J_{m,k}$ corresponding to the derivatives of the residual vector elements with respect to the first 5 elements of $\boldsymbol{x}$, *i.e.* α_{Ak}, ω_{DAk}, α_{B0}, ω_{DB0}, and τ_{B0}. The 'B' blocks are $N \times 2M$ matrices and are the derivatives of the residuals with respect to the remaining elements of $\boldsymbol{x}$, *i.e.* $X_{k(m)}$ and $Y_{k(m)}$ for $m = 1 \ldots M$. Many of the elements of the 'B' blocks will be zero, because the residuals in the m^{th} coherent integration only depend upon $X_{k(m)}$ and $Y_{k(m)}$—only 2 of the $2M$ columns in a given row contain non-zero entries. If we let $\tilde{\boldsymbol{x}}$ be the first 5 elements of $\boldsymbol{x}$ and $\breve{\boldsymbol{x}}$ be the two elements of $\boldsymbol{x}$ that have non-zero columns of $B_{Im,k}$ and $B_{Qm,k}$, then we can recognize that in the linearized problem,

$$J_{m,k}\boldsymbol{x} = \begin{bmatrix} A_{Im,k} \\ A_{Qm,k} \end{bmatrix} \tilde{\boldsymbol{x}} + \begin{bmatrix} \breve{B}_{Im,k} \\ \breve{B}_{Qm,k} \end{bmatrix} \breve{\boldsymbol{x}} = \boldsymbol{r}_{m,k} \tag{4.13}$$

where $\breve{B}_{Im,k}$ and $\breve{B}_{Qm,k}$ are the $N \times 2$ matrices that constitute the non-zero columns of $B_{Im,k}$ and $B_{Qm,k}$. If the orthonormal/upper-triangular factorization of the 'B' block matrix on the right-hand side of Eq. (4.13) yields the $N \times 2$ orthonormal factor Q and the 2×2 upper-triangular factor R, then we have

$$J_{m,k}\boldsymbol{x} = \begin{bmatrix} A_{Im,k} \\ A_{Qm,k} \end{bmatrix} \tilde{\boldsymbol{x}} + QR\breve{\boldsymbol{x}} = \boldsymbol{r}_{m,k} \tag{4.14}$$

After solving Eq. (4.14) for $\breve{\boldsymbol{x}}$ in terms of $\tilde{\boldsymbol{x}}$ and $\boldsymbol{r}_{m,k}$, the following definitions can be written.

$$\mathscr{A}_{m,k} = \begin{bmatrix} A_{Im,k} \\ A_{Qm,k} \end{bmatrix} - Q \left(Q^{\mathrm{T}} \begin{bmatrix} A_{Im,k} \\ A_{Qm,k} \end{bmatrix} \right) \tag{4.15}$$

and

$$\boldsymbol{\imath}_{m,k} = \boldsymbol{r}_{m,k} - Q\left(Q^{\mathrm{T}}\boldsymbol{r}_{m,k}\right) \tag{4.16}$$

Summing over all M intervals yields

$$\mathscr{A}_k^{\mathrm{T}}\mathscr{A}_k = \sum_{m=1}^{M} \mathscr{A}_{m,k}^{\mathrm{T}}\mathscr{A}_{m,k} \tag{4.17}$$

$$\mathscr{A}_k^{\mathrm{T}}\,\boldsymbol{\imath}_k = \sum_{m=1}^{M} \mathscr{A}_{m,k}^{\mathrm{T}}\,\boldsymbol{\imath}_{m,k} \tag{4.18}$$

We can now solve a much smaller optimization problem that approximates the full problem. The cost function to be minimized is

$$\tilde{f}(\tilde{\boldsymbol{x}}) = \frac{1}{2}\,\boldsymbol{\imath}_k^{\dagger}\,\boldsymbol{\imath}_k \tag{4.19}$$

The iteration increment is

$$\tilde{\boldsymbol{p}}_k = -(\mathscr{A}_k^{\mathrm{T}}\mathscr{A}_k)^{-1}\mathscr{A}_k^{\mathrm{T}}\,\boldsymbol{\imath}_k \tag{4.20}$$

and the $k+1^{\text{th}}$ estimate of $\tilde{\boldsymbol{x}}$ is

$$\tilde{\boldsymbol{x}}_{k+1} = \tilde{\boldsymbol{x}}_k + \kappa\tilde{\boldsymbol{p}}_k \tag{4.21}$$

Solving this approximate optimization problem is computationally tractable, whereas solving the exact optimization may not be, depending on the number of coherent integration intervals. Typically, fewer than ten iterations are required to converge on the locally optimal $\tilde{\boldsymbol{x}}$. Obviously, this algorithm takes longer to compute for cases in which more data are used. Optimizing a signal parameter vector for the SGH took much less time than optimizing a parameter vector for the omnidirectional antenna. For the SGH, this iteration takes about a minute on an Intel i7 quadcore processor running MATLAB. For the omnidirectional antenna, the iteration might run for 10 minutes. In all cases investigated, this algorithm converged to a locally optimal estimate of the state vector. In cases in which noise caused

there to be no clear peak, for instance, as in Figure 3.20(b), the fine acquisition algorithm finds the set of parameters that describe the location of the noise peak nearest the coarse acquisition result. Typically, this will result in an accurate estimate of carrier Doppler shift, but a poor estimate of subcarrier phase.

4.2 Signal Observables

Once the optimal vector of parameters is found, fine acquisition is complete, and observables may be extracted. An observable is a measurement of a signal property that is useful for orbit determination. Carrier phase, subcarrier phase, subcarrier code start time, subcarrier relative Doppler, and carrier Doppler shift are all signal properties that could be measured to provide an orbit determination algorithm its raw measurements.

A carrier phase measurement counts the number of carrier cycles that have been received since some reference time. The carrier phase may be difficult to track or poorly defined, depending on the number and duration of coherent integrations used per detection and depending on the accuracy with which the polynomial phase model in Eq. (2.10) represents the true phase. The carrier phase model allows steps changes in phase at the edges of coherent integration intervals, which make carrier phase ill-defined at those edges.

Subcarrier phase varies much more smoothly and is significantly easier to track. Measurements of subcarrier phase are potentially of interest. However, a single 48 kHz subcarrier cycle is more than six kilometers long. In order to make the measurement useful, the subcarrier phase would need to be measured to within a small fraction of a cycle, say one fiftieth or one hundredth. While this is possible at moderate C/N_0 (say, that attained using the SGH), it is difficult or impossible at low C/N_0 (that attained using the omnidirectional antenna).

Carrier-subcarrier relative Doppler shift is another potentially interesting measurement,

though it too would need to be measured to a precision difficult to attain at low C/N_0.

Carrier Doppler shift seems to be the "low-hanging fruit" among the observables. Carrier Doppler shift is readily measured to a fine enough precision to be of significant utility. In fact, an orbit determination algorithm is developed in Chapter 5 which uses carrier Doppler shift as its exclusive measurement. If other properly modeled measurements were included, the results of the orbit determination would only improve. The remainder of this chapter presents an analysis of the carrier Doppler shift measurement and its quality.

Carrier Doppler shift is the difference between the nominal transmission frequency and the actual received frequency at a particular reference time. Differences between the two frequencies arise due to spacecraft clock frequency error, receiver clock frequency error, and the effects of relative motion between the satellite and the receiver. The Doppler shift measurement D in radians/second at the center of the non-coherent detection/signal-acquisition interval is related to the signal detection parameters by

$$D = \omega_{DAk} + \alpha_{Ak} \left(MT_{coh}/2 \right) \tag{4.22}$$

Essentially, the Doppler shift measurement is the initial frequency offset propagated to the middle of the non-coherent interval using the Doppler rate estimate. Looked at in another way, the Doppler shift measurement is the mean carrier Doppler shift over the non-coherent interval given the choice of a quadratic model for the beat carrier phase made in Eq. (2.10), which translates into a linear model for the carrier Doppler shift. If a different polynomial order were used, then the correspondence between the observable that is the mid-point Doppler and the mean Doppler would no longer hold.

In addition to the raw observable, any orbit determination algorithm also needs an estimate of the uncertainty associated with each measurement. The uncertainty of the Doppler shift observable is measured by its variance. Three possible approaches to determining the variance of the carrier Doppler shift present themselves. First, an experiment could be car-

ried out much like the proof-of-concept experiment described in Chapter 3 in which the measured carrier Doppler statistics from the low-C/N_0 path were determined by comparison to the "truth" from the high-C/N_0 path. Second, a similar experiment could be carried out in simulation using identical truth parameters, but many different realizations of noise. Third, the variance can be calculated theoretically from the optimization described in Section 4.1. The drawbacks to the first two methods are that they would require a great deal of time. The third method is much faster. One simply calculates the Cramér-Rao Lower Bound (CRLB)[17] on the Doppler variance. The CRLB is the theoretical minimum covariance matrix of a parameter vector estimated from a least-squares problem. The actual covariance may be somewhat higher, but the CRLB will provide an estimate of at least the order of magnitude of the Doppler variance. The CRLB states that

$$\mathrm{E}\left[(\hat{\boldsymbol{x}}-\boldsymbol{x}_0)(\hat{\boldsymbol{x}}-\boldsymbol{x}_0)^{\mathrm{T}}\right] \geq F^{-1} \tag{4.23}$$

where F is the Fisher Information Matrix (FIM).

$$F = \mathrm{E}\left[\frac{\partial^2 \lambda(\boldsymbol{x})}{\partial \boldsymbol{x}^2}\right] \tag{4.24}$$

and

$$\frac{\partial^2 \lambda(\boldsymbol{x})}{\partial \boldsymbol{x}^2} = \frac{1}{\sigma^2}\frac{\partial^2 f(\boldsymbol{x})}{\partial \boldsymbol{x}^2} \tag{4.25}$$

with $f(\boldsymbol{x})$ as in Eq. (4.1). That is, the FIM is related to the approximate Hessian $J_k^{\mathrm{T}}J_k$ from Eq. (4.10), normalized by the variance of the residuals. That is, $F = \frac{1}{\sigma^2}J_k^{\mathrm{T}}J_k$, where the derivatives are evaluated at the final iterate of the estimate of the parameter vector. As discussed in Sec. 4.1, $J_k^{\mathrm{T}}J_k$ may be a large matrix, of which only the 5×5 upper left block is of interest. Conveniently, the corresponding block of the inverse of this matrix, which is the actual matrix of interest is exactly equal to the inverse of $\mathscr{A}_k^{\mathrm{T}}\mathscr{A}_k$ from Eq. (4.17), again with the derivatives evaluated at the final iterate of the parameter vector. Then $\tilde{F} = \frac{1}{\sigma^2}\mathscr{A}_k^{\mathrm{T}}\mathscr{A}_k$ and the minimum covariance matrix of the parameter vector $\mathrm{E}\left[(\tilde{\boldsymbol{x}}-\tilde{\boldsymbol{x}}_0)(\tilde{\boldsymbol{x}}-\tilde{\boldsymbol{x}}_0)^{\mathrm{T}}\right] \geq \tilde{F}^{-1}$.

When considering the variance of the Doppler shift measurement, only the first two diagonal elements of the covariance matrix are needed. In order to calculate the covariance,

first invert $\frac{1}{\sigma^2}\mathscr{A}_k^{\mathrm{T}}\mathscr{A}_k$.

$$P = \sigma^2 \left[\mathscr{A}_k^{\mathrm{T}}\mathscr{A}_k\right]^{-1} \tag{4.26}$$

P is the CRLB covariance matrix.

The variance sought is the variance of the Doppler shift at the mid-point of the detection interval, $\mathrm{E}\left[(\hat{\omega}_{Dmid} - \omega_{Dmid})(\hat{\omega}_{Dmid} - \omega_{Dmid})\right]$, where ω_{Dmid} is the true Doppler shift at the mid-point of the interval and $\hat{\omega}_{Dmid}$ is the estimate at the mid-point of the interval. As previously described, the estimate of the Doppler shift at the mid-point of the interval can be calculated from the signal detection parameters as $\hat{\omega}_{Dmid} = \hat{\omega}_{DAk} + \hat{\alpha}_{Ak} M T_{coh}/2$. If we say that the parameter estimates are equal to the truth plus some random measurement error, then we can write $\hat{\omega}_{DAk} = \omega_{DAk} + \epsilon_\omega$ and $\hat{\alpha}_{Ak} = \alpha_{DAk} + \epsilon_\alpha$. Now it is clear that the quantity sought is $\mathrm{E}\left[(\epsilon_\omega + \epsilon_\alpha M T_{coh}/2)(\epsilon_\omega + \epsilon_\alpha M T_{coh}/2)\right]$. Recognizing that $\mathrm{E}\left[\epsilon_\omega^2\right] = [P]_{1,1} = \sigma_{Dopp}^2$, $\mathrm{E}\left[\epsilon_\alpha \epsilon_\omega\right] = [P]_{1,2} = [P]_{2,1} = \sigma_{Dopp}\sigma_\alpha \rho_{\alpha Dopp}$, with $\rho_{\alpha Dopp}$ being the correlation coefficient, and $\mathrm{E}\left[\epsilon_\alpha^2\right] = [P]_{2,2} = \sigma_\alpha^2$, the Doppler variance at the mid-point of the interval can be calculated as

$$\sigma_{DoppMid}^2 = [P]_{1,1} + [P]_{1,2} M T_{coh} + [P]_{2,2} M^2 T_{coh}^2/4 \tag{4.27}$$

Interestingly and intuitively, there is always some anti-correlation between the Doppler error and the Doppler rate error, so the cross term always serves to reduce the estimate of the Doppler variance at the mid-point of the interval. This makes sense because the Doppler estimate at the mid-point is based on data both before and after that time point. Therefore, it should be more accurate than the estimate at the start of the interval, when the only available data lies after the time point of interest.

The Doppler shift variance has been calculated for three detection scenarios and the results are listed in Table 4.1. As a sanity check, the standard deviation calculated using the CRLB has been compared to the magnitude of the differences between the "true" Doppler shifts as computed using high-SNR data from the proof-of-concept experiment and those computed from minimally aided detections using the low-SNR data. Reasonable results, like

those shown in the table, were attained for all low-SNR detections. The last line of the table shows that one detection resulted in Doppler error of less than the CRLB. This can happen by chance, and in fact, it should occasionally happen. What is important is that the three values in the fourth column are reasonable given the CRLB values in the last column; none of these values is many times larger than the corresponding CRLB value. Also, the CRLB values vary with acquisition parameters in the first 3 columns as one would expect. In general, longer detection intervals are more sensitive to errors in Doppler shift. So, all other things being equal, Doppler shifts obtained from longer detections are more accurate than those obtained from shorter detections. Also, all other things being equal, Doppler shifts obtained from higher-C/N_0 data are more accurate than those obtained from lower-C/N_0 data.

Table 4.1 – Influence of acquisition parameters on standard deviation of Doppler shift measurement error.

C/N_0 dB-Hz	T_{coh} s	M –	Doppler Error* Hz	CRLB† Hz
10.956	0.125	240	0.554	0.406
31.956	0.25	6	0.048	0.023
31.956	0.125	12	0.010	0.124

*Difference between high- and low-SNR detections at midinterval.
†Square root of CRLB quantities projected to mid-interval as in Eq. (4.27).

CHAPTER 5

ORBIT DETERMINATION

The goal of this work is to lay the foundation for a low-cost orbit determination system that can support satellites during orbit raising, that is, the stage of the satellite lifetime from after decoupling with the launch vehicle until insertion into the final orbital slot and handoff to the customer. Thus far, it has been demonstrated that the satellite telemetry signals can be reasonably detected and what level of quality can be expected from the derived carrier Doppler shift observables of the signals. The final analysis necessary is to show that using these observables, reasonable levels of orbit determination accuracy can be achieved with a reasonable number of ground stations.

This chapter presents an orbit and measurement simulation, a simple orbit determination filter, and the orbit determination results of applying the filter to the simulated measurements. The models presented in this chapter are not intended to be high-fidelity models. They are intended to be rough models that provide a level of fidelity sufficient to inform decisions regarding the overall design and disposition of future low-cost ground stations. It would have been desirable to include some orbit determination results using on-orbit data. Unfortunately, sufficient data are not available. The proof-of-concept experiment was designed to provide data for signal detection only. All the results presented in this chapter use simulated orbits and measurements with realistic levels of process noise and measurement noise, based on the results from the previous chapter.

The next section provides a summary of the orbit and clock simulation. After that, the extended square-root information orbit determination filter is described. The cases simulated, including different ground station numbers and locations, are detailed in the following section. The final section of the chapter gives results of the orbit determination simulations.

5.1 Orbit and Clock Model

The satellite orbit and its clock were modeled in discrete time *via* a non-linear dynamic state transition function. The simulation and the orbit determination filter used the same dynamic state transition function,

$$\boldsymbol{x}_{k+1} = \boldsymbol{f}_k(\boldsymbol{x}_k, \boldsymbol{w}_k) \tag{5.1}$$

where $\boldsymbol{x}_k$ is the state vector and $\boldsymbol{w}_k$ is a random process noise vector, k is the sample index which takes on values $k = 0, 1, 2, \ldots$, and $\boldsymbol{f}_k$ is a vector-valued function of the two input vectors $\boldsymbol{x}_k$ and $\boldsymbol{w}_k$. It models how the physical system evolves with time. In this case, the physical system is the satellite's orbit and its clock. Note that some notational reuse has occurred, and that the $\boldsymbol{f}_k(\boldsymbol{x}_k, \boldsymbol{w}_k)$ of this chapter is a completely different quantity from the $f(\boldsymbol{x})$ from the previous chapter.

The state vector consists of 17 elements. Three state vector elements describe the satellite's Cartesian position vector in an Earth-centered inertial reference frame (ECIF), three describe Cartesian velocity in the same frame and measured with respect to that frame, two describe the spacecraft clock state, and the remaining nine state vector elements describe additional states referred to in the literature as empirical acceleration parameters [18, 19, 20].

The process noise vector consists of 14 elements. Three provide random velocity increments, *i.e.*, impulsive acceleration disturbances, each time step. Two drive the spacecraft clock model. Three drive the random walk of the "constant" along-track, altitude, and cross-track velocity disturbances. And three each drive the random walks of the coefficients of the cosine and sine once-per-orbit velocity disturbances. The rest of this section describes in detail the dynamics and statistics of each element of the state and process noise vectors.

The simulated dynamics of the satellite position and velocity are purely Keplerian, except for the addition of white noise and the effects of the empirical accelerations on the velocity

states. The orbit determination filter's dynamics model uses the same dynamics model as the orbit simulation. They are propagated through time *via* a transformation from Cartesian to equinoctial coordinates, a time propagation in equinoctial coordinates in which only the mean longitude[1] changes[21], and lastly a transformation back to Cartesian. The transformation to and from equinoctial elements is used purely as a convenience. In equinoctial coordinates, propagation of the orbit is trivial, whereas in Cartesian coordinates, orbit propagation is somewhat more involved. At the end of each propagation interval, the Keplerian satellite velocity states are modified by the addition of the effects of the empirical accelerations and by the addition of white noise. The velocity modification can be expressed as

$$\boldsymbol{v}_{k+1} = \tilde{\boldsymbol{v}}_{k+1}(\boldsymbol{r}_k, \boldsymbol{v}_k) + E(\boldsymbol{r}_k, \boldsymbol{v}_k)\left(\boldsymbol{a}_{0,k+1} + \boldsymbol{a}_{1,k+1}\cos\lambda_{k+1} + \boldsymbol{a}_{2,k+1}\sin\lambda_{k+1} + \boldsymbol{w}_v\right) \tag{5.2}$$

where the components of $\boldsymbol{v}_{k+1}$ are the fourth through sixth elements of $\boldsymbol{x}_{k+1}$, the satellite velocity, $\tilde{\boldsymbol{v}}_{k+1}(\boldsymbol{r}_k, \boldsymbol{v}_k)$ is the Keplerian velocity at time t_{k+1} based on the position and velocity at time t_k, $E(\boldsymbol{r}_k, \boldsymbol{v}_k)$ is the matrix that transforms from local level coordinates (along-track, altitude, and cross-track) to ECIF coordinates, $\boldsymbol{a}_{0,k+1}$ is the constant velocity increment in local level coordinates due to the effects of the empirical accelerations which applies at time t_{k+1}, $\boldsymbol{a}_{1,k+1}$ and $\boldsymbol{a}_{2,k+1}$ are the velocity increments due to once-per-orbital-revolution acceleration disturbances which apply at time t_{k+1}, λ_{k+1} is the mean longitude, and $\boldsymbol{w}_v$ is a vector of zero-mean, white, Gaussian process noise.

The standard deviation of the Gaussian distribution from which the elements of the process noise vector $\boldsymbol{w}_v$ are drawn is set such that the resulting equivalent accelerations are roughly equal to, or greater than, the magnitudes of unmodeled accelerations that a real satellite in the simulated orbit would experience [22]. For instance, in GTO, accelerations down to approximately 1×10^{-7} to 1×10^{-10} m/s^2 would be modeled over most of the orbit. Over a 600-second interval (the discrete time step of the simulation and filter), this would result in a velocity increment of, say, 600×10^{-8} m/s for acceleration disturbances of

[1] In terms of traditional orbital elements, the mean longitude is the sum of the mean anomaly, the argument of perigee, and the longitude of ascending node, $\lambda = M + \omega + \Omega$

1×10^{-8} m/s^2. Accordingly, the white noise driving the velocity evolution of the satellite is set to have standard deviation of 600×10^{-8} m/s for the GTO orbit.

The effects of the empirical accelerations are cast as velocity disturbances that occur at the end of each propagation interval. The empirical accelerations are simulated and added to the filter's dynamic model in order to accommodate the effects of unmodeled or mismodeled perturbations to the satellite's orbit. These accelerations would be due to, for instance, error or mismatch in the atmospheric density model, the spacecraft drag model, the solar radiation pressure model, the Earth albedo model, *etc.* In some of the literature, twice-per-orbit empirical acceleration disturbances are also modeled [20, 23]. The nine empirical acceleration effect coefficients, $\boldsymbol{a}_{k+1} = [\boldsymbol{a}_{0,k+1}^{\mathrm{T}}, \boldsymbol{a}_{1,k+1}^{\mathrm{T}}, \boldsymbol{a}_{2,k+1}^{\mathrm{T}}]^{\mathrm{T}}$, comprise the last nine elements of the state vector. These elements evolve according to a random walk model of the form

$$\boldsymbol{a}_{k+1} = \boldsymbol{a}_k + \boldsymbol{w}_a \tag{5.3}$$

where $\boldsymbol{a}_k$ and $\boldsymbol{a}_{k+1}$ are the vectors of empirical acceleration effect coefficients which apply at the subscripted time steps and $\boldsymbol{w}_a$ is a nine-element zero-mean, white, Gaussian process noise vector. The initial values of the empirical acceleration effect coefficients are drawn from a zero-mean Gaussian distribution with standard deviation the same as the standard deviation of the elements of $\boldsymbol{w}_v$. That is, the empirical accelerations are modeled as being about the order of magnitude of the Gaussian white-noise unmodeled or mismodeled accelerations. The process noise vector elements, $\boldsymbol{w}_a$, that drive the random walk evolution of the empirical acceleration effect coefficients are drawn from a distribution with standard deviation that is two orders of magnitude lower than the standard deviation of the initial values of the coefficients. For reference, it takes 630 orbits for the empirical process noise coefficients to double in standard deviation in the absence of Kalman filter measurement updates[2]. So, the

[2]After N steps, the variance of a parameter undergoing a discrete random walk with steps of variance σ^2 is $\sigma_p^2 = N\sigma^2$. In this case, the standard deviation of interest is double the original, that is $\sigma_p = 2\sigma_0$, and the empirical process noise coefficients were drawn from an original distribution with 100 times more standard deviation than the standard deviation of the steps perturbing them, that is $\sigma_0 = 100\sigma$, so $4\sigma_0^2 = N(1 \times 10^{-2}\sigma_0)^2$. N is easily found to be 40000. With orbital period of 38096.8 s and time steps of 600 s, the number of orbits until the standard deviation of the coefficients doubles is 629.9.

coefficients walk randomly, but not rapidly.

The clock model consists of a standard two-state Markov model [24],

$$\begin{bmatrix} \delta t_{Ak+1} \\ \dot{\delta} t_{Ak+1} \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \delta t_{Ak} \\ \dot{\delta} t_{Ak} \end{bmatrix} + \begin{bmatrix} \Delta t & 0 \\ 0 & \Delta t \end{bmatrix} \begin{bmatrix} w_t \\ w_f \end{bmatrix} \tag{5.4}$$

in which the satellite clock rate error $\dot{\delta} t_{Ak+1}$ evolves according to a random walk and the satellite clock error is the integral of the rate error plus an additional random walk. The Markov model is driven by the process noise vector elements w_t and w_f which are white, zero-mean, and independent with equivalent continuous-time power spectral densities S_f and S_g respectively. S_f and S_g are related to the Allan variance parameters by $S_f \approx h_0/2$ and $S_g \approx 2\pi^2 h_{-2}$.

Two types of orbit were simulated. The first is a geostationary transfer orbit with periapsis altitude of 400 km ($r_p = 6771$ km) and apoapsis altitude of 35800 km ($r_a = 42171$ km), which is slightly higher than synchronous orbit. The second orbit simulated is an intermediate orbit between the initial GTO and GEO with periapsis altitude of 17700 km ($r_p = 24071$ km) and apoapsis altitude of 35800 km ($r_a = 42171$ km). Both orbits are Keplerian, with the exception of the addition of the random velocity increments and the empirical accelerations.

For the GTO orbit, the addition of process noise and empirical accelerations results in a maximum drift in the along-track direction of 7.1 km, of 2.8 km in the altitude direction, and of 10.6 km in the cross-track direction over the simulated two-week period for one realization of process noise. For the intermediate orbit, the addition of process noise and empirical accelerations results in a maximum drift in the along-track direction of 3.4 km, of 0.9 km in the altitude direction, and of 4.3 km in the cross-track direction over the simulated two-week period for one realization of the process noise.

Measurements, including random measurement noise, were also simulated for both orbits

and for all of the ground stations listed in Table 5.1. The standard deviation of the carrier Doppler shift measurement noise corresponds roughly to the levels determined in Sec. 4.2, either 0.5 Hz or 0.1 Hz depending on the particular simulation.

5.2 Extended Square-Root Information Filter

In order to determine roughly how well real-time orbit determination would function using these signals and the resulting observables, an extended square-root information filter (ESRIF) was applied to the problem. The ESRIF has been extensively studied and is well-understood [25, 26]. In fact, such filters have been applied to orbit determination in several instances [27, 28, 29]. It is a suboptimal filter due to its use of linear approximations, but it is simple to implement and works well enough for the relatively mild nonlinearities encountered over the likely level of filter state uncertainty in GEO or GTO and with an average of several ground station contacts per orbit. The remainder of this section describes the iterative filtering process used for this study.

An extended square root information filter uses a linearized dynamics function. The non-linear dynamics function in Eq. (5.1) can be linearized by taking its derivatives with respect to the state and the process noise, which results in the Jacobian matrices F_k and Γ_k respectively. Explicitly,

$$F_k = \left.\frac{\partial \boldsymbol{f}_k(\boldsymbol{x}_k, \boldsymbol{w}_k)}{\partial \boldsymbol{x}_k}\right|_{\boldsymbol{x}_k=\hat{\boldsymbol{x}}_k, \boldsymbol{w}_k=\boldsymbol{0}} \tag{5.5}$$

$$\Gamma_k = \left.\frac{\partial \boldsymbol{f}_k(\boldsymbol{x}_k, \boldsymbol{w}_k)}{\partial \boldsymbol{w}_k}\right|_{\boldsymbol{x}_k=\hat{\boldsymbol{x}}_k, \boldsymbol{w}_k=\boldsymbol{0}} \tag{5.6}$$

Then the linearized *a priori* state estimate error can be written

$$\boldsymbol{x}_{k+1} - \bar{\boldsymbol{x}}_{k+1} \approx F_k \left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k\right) + \Gamma_k \boldsymbol{w}_k \tag{5.7}$$

where $\boldsymbol{x}_k$ is the true state vector at index k, $\bar{\boldsymbol{x}}_{k+1} = \boldsymbol{f}_k(\hat{\boldsymbol{x}}_k, \boldsymbol{0})$ is the *a priori* estimate of the state vector at index $k+1$, that is, it is the state estimate at sample $k+1$ conditioned only

on measurements up through sample k, and $\hat{\boldsymbol{x}}_k$ is the *a posteriori* estimate of the state at sample k, that is, it is the state estimate at sample k conditioned on measurements right up through sample k.

In addition to the state transition function, the filter requires a measurement model. A simple model of the Doppler shift observable in Hz, excluding ionospheric effects, can be written

$$D = \frac{-1}{\lambda_A}\left(a_x(\dot{X}_A - \dot{X}_R) + a_y(\dot{Y}_A - \dot{Y}_R) + a_z(\dot{Z}_A - \dot{Z}_R)\right) + f_A\dot{\delta t}_A + n_D \tag{5.8}$$

In this model, λ_A is the nominal carrier wavelength in meters, a_i for $i \in \{x, y, z\}$ is the i component in Earth-centered inertially-fixed Cartesian coordinates of the unit vector which points from the receiver to the satellite,

$$\begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} = \frac{\boldsymbol{r}_A - \boldsymbol{r}_R}{\sqrt{(\boldsymbol{r}_A - \boldsymbol{r}_R)^{\mathrm{T}}(\boldsymbol{r}_A - \boldsymbol{r}_R)}} \tag{5.9}$$

where the satellite's position is $\boldsymbol{r}_A = [X_A,\ Y_A,\ Z_A]^{\mathrm{T}}$ in ECIF coordinates and its velocity is $\boldsymbol{v}_A = [\dot{X}_A,\ \dot{Y}_A,\ \dot{Z}_A]^{\mathrm{T}}$, the receiver's position expressed in ECIF coordinates is $\boldsymbol{r}_R = [X_R,\ Y_R,\ Z_R]^{\mathrm{T}}$ and its velocity is $\boldsymbol{v}_R = [\dot{X}_R,\ \dot{Y}_R,\ \dot{Z}_R]^{\mathrm{T}}$, f_A is the nominal carrier frequency, $\dot{\delta t}_A$ is the spacecraft clock error rate in s/s, as in Eq. (5.4), and n_D is zero-mean, white, Gaussian measurement noise. The Doppler shift D is related to the signal detection parameters by $D = \omega_{DAk} + \alpha_{Ak}\,(2MT_{coh}/2)$. This simple model does not take into account propagation delay. All quantities apply at the same instant. Nor does it account for the additional delay due to the neutral atmosphere or carrier phase advance due to the ionosphere. Also, it neglects the possible effects of receiver clock error on the assumption that this contribution will be removed by use of GPS disciplining of the receiver clock. This is sufficient to model the relationship between what the satellite is doing and what the ground stations observe for a rough analysis of orbit determination using this observable derived

from low-cost ground stations. This model would not be adequate for a deployed orbit determination system, because additional higher-order effects are needed in order to achieve the best possible accuracy, effects such as the light-time delay between the transmission from the satellite and the time of reception on the ground.

The ESRIF's measurement model takes the standard form

$$\boldsymbol{y}_{k+1} = \boldsymbol{h}_{k+1}(\boldsymbol{x}_{k+1}) + \boldsymbol{\nu}_{k+1} \tag{5.10}$$

where the vector of normalized measurements available at epoch $k+1$ is modeled as the sum of a non-linear vector-valued function, $\boldsymbol{h}_{k+1}(\boldsymbol{x}_{k+1})$ and a zero-mean, unit variance, white, Gaussian noise vector $\boldsymbol{\nu}_{k+1}$. The non-linear measurement model function $\boldsymbol{h}_{k+1}(\boldsymbol{x}_{k+1})$ captures the relationship of the measurements to the spacecraft state vector $\boldsymbol{x}_{k+1}$. The Doppler shift model of Eq. (5.8) represents one element of the output of $\boldsymbol{h}_{k+1}(\boldsymbol{x}_{k+1})$. The conditioning on the variance of the measurement noise is enforced through normalization of the actual measurements. That is, $\boldsymbol{\nu}_{k+1} = \boldsymbol{n}_{Dk+1}/\sigma_D$, where $\boldsymbol{n}_{Dk+1}$ is the vector of measurement noise elements, one element of which is shown in Eq. (5.8) above, and σ_D^2 is the Doppler measurement variance.

The ESRIF also requires a linearization of the measurement model. If the Jacobian matrix of the measurement vector with respect to the state vector is called H_{xk+1}, then the linearization is

$$\Delta\boldsymbol{y}_{k+1} = H_{xk+1}\left(\boldsymbol{x}_{k+1} - \bar{\boldsymbol{x}}_{k+1}\right) + \nu_{k+1} \tag{5.11}$$

where $\Delta\boldsymbol{y}_{k+1}$ is the difference between the actual measurements and those expected based on the measurement model with $\bar{\boldsymbol{x}}_{k+1}$ as input, $\Delta\boldsymbol{y}_{k+1} = \boldsymbol{y}_{k+1} - \boldsymbol{h}_{k+1}(\bar{\boldsymbol{x}}_{k+1})$.

With these preliminaries taken care of, the iterative filtering process can be simply ex-

plained. First, the *a priori* square-root information equations are written

$$\begin{bmatrix} R_{wwk} & 0 \\ 0 & \hat{R}_{xxk} \end{bmatrix} \begin{bmatrix} \boldsymbol{w}_k \\ \boldsymbol{x}_k - \hat{\boldsymbol{x}}_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} \boldsymbol{\nu}_{wk} \\ \boldsymbol{\nu}_{xk} \end{bmatrix} \tag{5.12}$$

In the above equations, R_{wwk} is the *a priori* process noise square-root information matrix—the inverse square root of the process noise covariance matrix, $\hat{R}_{xxk}$ is the *a posteriori* satellite state square-root information matrix, $\boldsymbol{w}_k$ is the process noise vector acting from k to $k+1$, and $\boldsymbol{\nu}_{wk}$ and $\boldsymbol{\nu}_{xk}$ are unknown zero-mean, idnetity-covariance, uncorrelated white-noise vectors. These equations essentially state that the process noise and the state error properly normalized have unity variance. These equations involve $\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k$, but we would like them to involve $\boldsymbol{x}_{k+1} - \bar{\boldsymbol{x}}_{k+1}$. This can be accomplished by solving Eq. (5.7) for $\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k$ and substituting the result into Eq. (5.12). At the same time, the linearized measurement equation can be appended to the system of equations.

$$\begin{bmatrix} R_{wwk} & 0 \\ -\hat{R}_{xxk}F_k^{-1}\Gamma_k & \hat{R}_{xxk}F_k^{-1} \\ 0 & H_{xk+1} \end{bmatrix} \begin{bmatrix} \boldsymbol{w}_k \\ \boldsymbol{x}_{k+1} - \bar{\boldsymbol{x}}_{k+1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \Delta\boldsymbol{y}_{k+1} \end{bmatrix} - \begin{bmatrix} \boldsymbol{\nu}_{wk} \\ \boldsymbol{\nu}_{xk} \\ \boldsymbol{\nu}_{k+1} \end{bmatrix} \tag{5.13}$$

The final phase in the filtering step is to employ an orthonormal/upper-triangular (QR) factorization [26] in order to solve the system of equations. The large block matrix will be factored into an orthonormal matrix and an upper triangular matrix. If the entire system of equations is multiplied by the transpose of the orthonormal factor, the result is

$$\begin{bmatrix} \hat{R}_{wwk} & \hat{R}_{wxk} \\ 0 & \hat{R}_{xxk+1} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{w}_k \\ \boldsymbol{x}_{k+1} - \bar{\boldsymbol{x}}_{k+1} \end{bmatrix} = \begin{bmatrix} \Delta\boldsymbol{z}_{wk} \\ \Delta\boldsymbol{z}_{xk+1} \\ \Delta\boldsymbol{z}_{resk+1} \end{bmatrix} - \begin{bmatrix} \hat{\boldsymbol{\nu}}_{wk} \\ \hat{\boldsymbol{\nu}}_{xk+1} \\ \hat{\boldsymbol{\nu}}_{k+1} \end{bmatrix} \tag{5.14}$$

If the middle of these three equations is solved for $\boldsymbol{x}_{k+1}$, the result is

$$\boldsymbol{x}_{k+1} = \bar{\boldsymbol{x}}_{k+1} + \hat{R}_{xxk+1}^{-1}\left(\Delta\boldsymbol{z}_{xk+1} - \hat{\boldsymbol{\nu}}_{xk+1}\right) \tag{5.15}$$

The expected value of this equation is:

$$\hat{\boldsymbol{x}}_{k+1} = \bar{\boldsymbol{x}}_{k+1} + \hat{R}_{xxk+1}^{-1}\Delta\boldsymbol{z}_{xk+1} \tag{5.16}$$

The filtering process proceeds to the next time step by passing forward $\hat{\boldsymbol{x}}_{k+1}$ and $\hat{R}_{xxk+1}$ and retrieving the next set of measurements.

In order to begin the filtering process, an initial estimate of the spacecraft state is needed. For both orbits, the filter was initialized with a random initial position estimate with RMS error of 100 m in each direction and a random velocity estimate with RMS error of 0.030 m/s in each direction. These levels of uncertainty correspond to those available from the launch vehicle operator. An operational orbit determination system would get its initial spacecraft state estimate in a handoff from the launch vehicle, or in a handoff from the traditional orbit determination system. The 2-week Kalman filter simulation scenarios are long enough to make their average and final accuracies largely insensitive to the choices of initial uncertainty defined here. Therefore, any averaged or final results will be indicative of the possible system performance, independent of the validity of the assumed filter initialization error standard deviations.

5.3 Ground Station Placement

Seventeen ground stations were used in the orbit determination simulations. These ground stations are located around the world, at the locations listed in Table 5.1. They are placed on the surface of a sphere of the mean radius of the Earth, 6371 km.

In the simulations, the ground station location is known exactly. In a real system, the actual location of the phase center of the receiving antenna is not known perfectly. This uncertainty introduces biases into the measurements which affect the final quality of the orbit determination. Depending on the required level of orbit determination accuracy, it may be necessary to find the actual incident-angle-dependent phase center location to within centimeters or millimeters [30, 31, 32, 33].

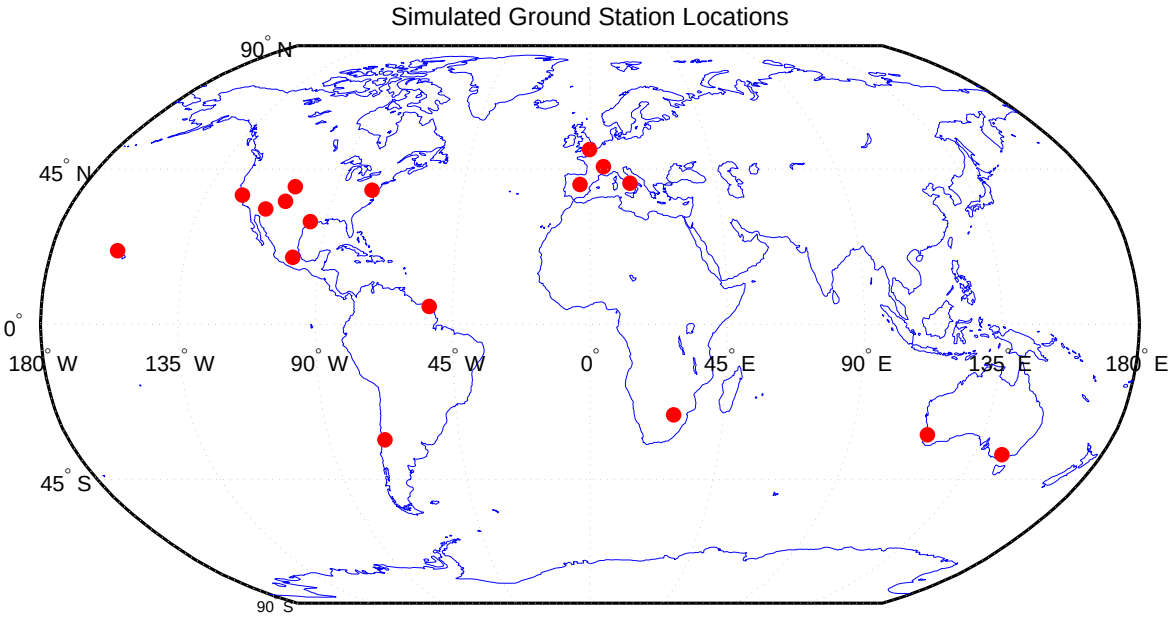
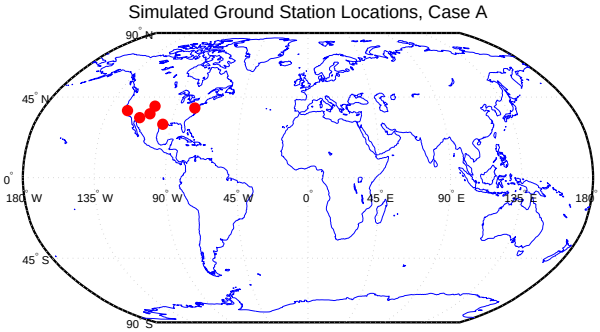
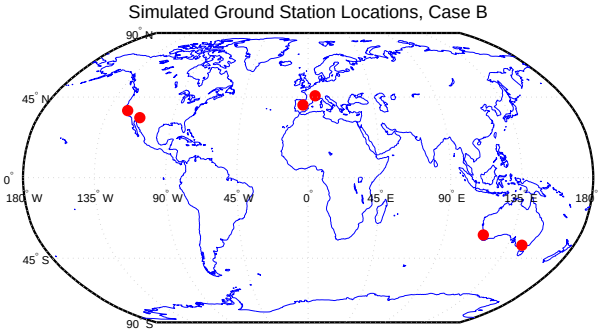


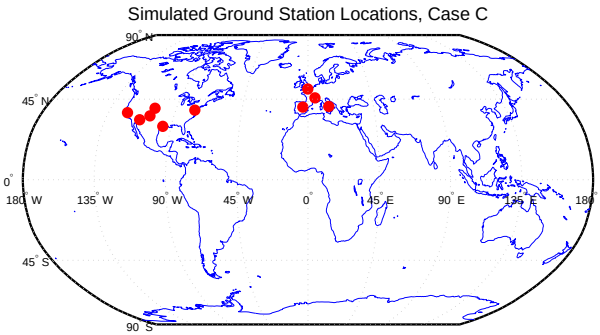
Figure 5.1 – Map of locations of simulated ground stations.



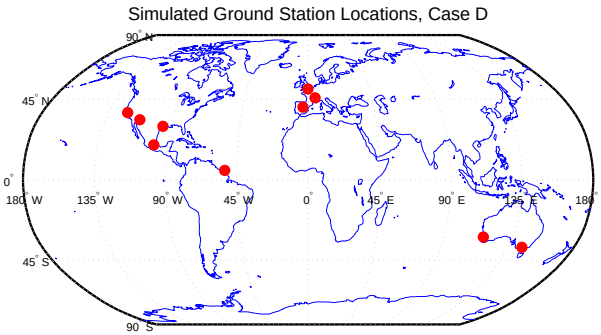
(a) Case A ground stations.



(b) Case B ground stations.



(c) Case C ground stations.



(d) Case D ground stations.

Figure 5.2 – Maps of ground stations used in cases A–D.

Table 5.1 – Locations of ground stations used in the orbit determination simulation.

Sta. Num.	Location	N. Lat. (°)	E. Long. (°)
1	Palo Alto, CA	37.426233	-122.103400
2	Phoenix, AZ	33.435627	-112.019492
3	Denver, CO	39.860465	-104.673494
4	Santa Fe, NM	35.618256	-106.085131
5	Houston, TX	29.756817	-95.362633
6	Arlington, VA	38.846365	-77.043147
7	Lyon, France	45.731930	4.940435
8	Madrid, Spain	40.519436	-3.568300
9	Brighton, UK	50.831945	-0.140967
10	Naples, Italy	40.884310	14.290204
11	Mexico City, Mexico	19.443249	-99.051462
12	Kourou, French Guiana	5.219365	-52.763415
13	Perth, Australia	-31.940406	115.954839
14	Melbourne, Australia	-37.680476	144.856209
15	Honolulu, HI	21.321094	-157.933394
16	Johannesburg, South Africa	-26.146197	28.243538
17	Santiago, Chile	-33.394647	-70.798928

Five ground station cases are considered for each of the two simulated orbits. The cases, lettered A–E, are listed in Table 5.2. In Case A there are six ground stations located in the United States. The locations of these ground stations are shown in Fig. 5.2(a). In Case B, there are three pairs of ground stations scattered across the globe—one pair located in the Western US, one pair located in Europe, and the last pair located in Australia. Fig. 5.2(b) is a map showing the locations of the Case B ground stations. Case C uses 10 ground stations in the US and Europe. They are shown in Fig. 5.2(c). Case D also uses 10 ground stations, but they are more scattered around the world. They are shown in Fig. 5.2(d). Case E uses all 17 ground stations listed in Table 5.1 and shown in Fig. 5.1. Some of the ground stations were placed near existing Earth stations. The locations of the other ground stations are in no way special or significant, except that they provide geographic diversity and are located roughly in groups of two or more. The ground stations are grouped because if the spacecraft is observed with only one ground station at a time, the spacecraft clock error is inseparable from the orbit-induced Doppler shift. Observing the spacecraft with more than one ground

Table 5.2 – Ground stations used for each of the 5 cases, listed by station number.

Case	Stations Used
A	1–6
B	1, 2, 7, 8, 13, 14
C	1–10
D	1, 2, 5, 7–9, 11–14
E	1–17

station provides the orbit determination algorithm with useful orbit information from the Doppler shift observables.

For each of cases A–E, the appropriate simulated measurements were made available to the filter. In order for a measurement to be used by the filter, there had to be two or more ground stations in view at a particular measurement epoch, the satellite had to be more than ten degrees above the horizon, and the satellite had to be above 6400 km in altitude. This altitude constraint (roughly one Earth radius) is arbitrary and serves only to limit the angular rate of the satellite across the sky from the point-of-view of the ground stations. This reduces the importance of the changing dynamic situation of the spacecraft during the time it takes to make a measurement. In the low-C/N_0 regime, measurements take 30 or more seconds to make. Near perigee of a GTO orbit, the spacecraft may sweep through 40° in 30 seconds from the point of view of a ground station directly below the spacecraft, whereas it will sweep through less than 2° in 30 seconds when at an altitude of one Earth radius. While it would certainly be possible to slew a small antenna this quickly, maintaining dynamic pointing might increase the complexity of the system. With this altitude constraint, the satellite can be assumed to remain in the main beam of a statically pointed receiving antenna during the entire measurement duration.

5.4 Results

The results of several filter runs are presented in Table 5.3. All filter runs simulated 14 days with measurements every 10 minutes while the satellite was in view. Note that measurements may actually be available much more often than every ten minutes. As mentioned above, a measurement can be made using an omnidirectional antenna in roughly 30 seconds. Using the higher-gain SGH antenna, it may take as little as 1.5 seconds of data to make an independent carrier Doppler shift measurement. As a result, the filter used only 1.5 s–30 s-worth of data every 10 minutes. If measurements are made available to the filter more often than every ten minutes, filter performance might be better. The left-most column of the table indicates which set of ground stations was used, which orbit was simulated, and which level of measurement noise was used. The remaining columns list the RMS position estimate error magnitude, the peak position estimate error magnitude, the filter's estimate of position estimate standard deviation, the RMS velocity estimate error magnitude, and the peak velocity estimate error magnitude. The statistics are computed over days 8–14 of all filter runs. The entries in the table are arranged in groups of five lines each. Each group of five lines uses the same orbit (indicated by the numeral in the case name—'1' for the GTO orbit and '2' for the intermediate orbit) and the same carrier Doppler shift measurement noise standard deviation (indicated by the case of the letter in the designation—uppercase for 0.5 Hz standard deviation measurement noise, lowercase for 0.1 Hz standard deviation), but a different set of ground stations. Each of the 20 lines of the table represents the results of 100 filter runs with independent realizations of process and measurement noise. That is, the peak numbers, for instance, are the maximum absolute difference between the filter's estimate of position and the true position in 100 independent filter runs over the last 7 days of each run. Likewise, each table entry for RMS position error is the square root of the mean of the squares of the magnitudes of the differences between the filter's estimate of position and the true position in the same 100 filter runs over the same span of 7 days as used in

finding the corresponding peak number.

Table 5.3 – Orbit determination performance for 20 scenarios

	Position Error, m			Velocity Error, cm/s	
Case	RMS	Peak	Filter Std	RMS	Peak
A1	2793.98	32709.94	271.31	70.995	2799.41
B1	287.77	8196.06	168.03	11.572	689.16
C1	1593.74	10788.60	184.53	29.646	916.18
D1	384.90	3962.37	85.20	5.238	121.65
E1	42.38	268.98	42.25	1.135	23.32
a1	1132.47	9889.82	87.50	22.450	867.88
b1	78.75	1739.30	45.77	2.313	152.05
c1	681.04	7912.01	57.32	9.728	283.32
d1	58.21	838.28	25.39	1.142	26.53
e1	12.63	84.39	13.24	0.343	7.33
A2	305.04	4635.67	236.22	3.351	66.75
B2	174.37	793.47	181.74	1.916	11.43
C2	167.96	973.92	176.77	1.773	14.57
D2	105.71	429.88	112.97	1.139	6.30
E2	60.65	274.99	70.94	0.641	3.77
a2	83.09	523.93	85.76	0.885	7.66
b2	49.36	255.24	55.02	0.539	3.23
c2	56.58	363.12	59.20	0.605	5.28
d2	32.90	138.07	34.75	0.349	2.08
e2	19.55	105.63	21.31	0.201	1.46

The results shown in the table indicate that reasonable levels of orbit determination accuracy should be achievable using a modestly-sized network of receivers on the ground. RMS position estimate errors of less than 1.5 kilometers can be achieved with only six receivers in the contiguous United States. RMS position estimate errors of less than 50 meters can be achieved with only 17 receivers placed in easily-accessible locations around the world.

Comparing 'B' cases and 'C' cases is interesting. The 'B' cases use only six receivers, while the 'C' cases use ten. However, the 'B' cases tend to out-perform the 'C' cases. This is because there is greater geographical diversity of ground stations in the 'B' cases. They use one pair of ground stations in the US, one pair in Europe, and one pair in Australia. The 'C' cases use ground stations located only in the US and Europe. Adding ground

stations without significantly increasing geographic diversity is only marginally beneficial. Considerations such as this are important in choosing the locations of actual ground stations.

In addition to the tabulated results, some plots of position estimation error versus time are included in Figures 5.3–5.6. The figures show the history of position estimation error components and magnitude for one example from each of cases A1, E1, A2, and E2. The figures also show, in the (b) panels, a quantity computed from the filter's square root information matrix. It is genearted by computing the square root of the trace of the 3-by-3 sub-matrix associated with the position states of the filter's computed estimation error covariance matrix. In general, the orbit determination performance is improved significantly if higher-quality measurements are used. Only cases using the lower-quality measurements are depicted in the figures, but results for higher-quality measurements are shown in the table.

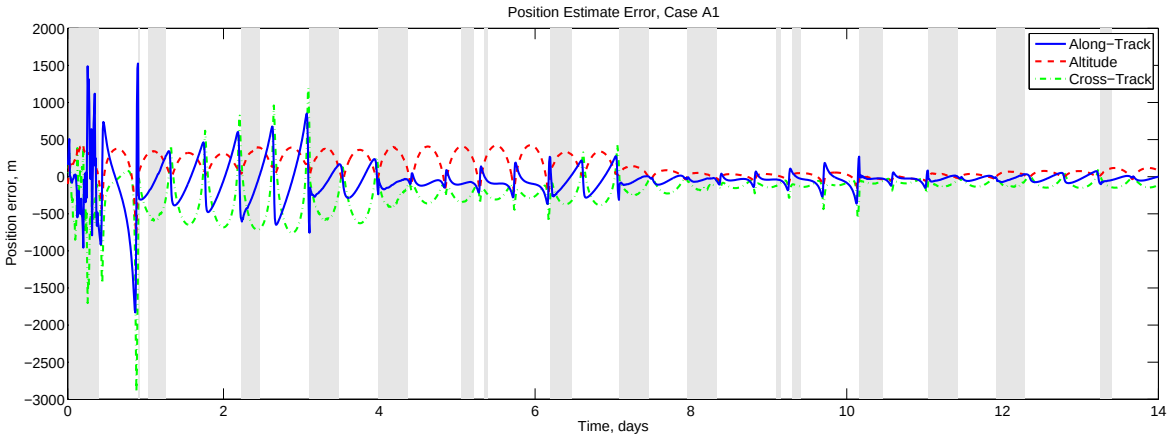
Despite relatively sparse and brief periods of satellite visibility, case A1 performs fairly well. That is, this set of only six low-cost ground stations produces orbit estimates with errors that do not grow without bound, and in fact might even be useful for some purposes. For instance, an error of 3 km gives the satellite location to less than half a degree of azimuth or elevation. This is sufficient precision to point a 30-inch satellite dish. Though not a part of the concept of operation of the system under investigation, a possible concept of operations might allow the orbit estimates to coast at this coarse level of precision, then periodically update the orbit estimate with higher-quality measurements to bring the uncertainty down during critical times, perhaps during maneuvers. In the example shown in the figure, there are gaps of nearly a full day during which no measurements are available to update the filter. During the data gaps, the filter's position estimate drifts away from the truth only moderately. Though, from the table, it is clear that in at least one of the simulated cases, the filter's estimate drifted away by nearly 33 km. The position error snaps down to a lower level at the onset of each grey region, as expected because of the ensuing filter measurement updates.

The example Case E1 shown in Figure 5.4 used the same realizations of process noise and measurement noise as the example Case A1 shown in Figure 5.3. The differences are purely due to which ground stations observed the satellite. Because Case E1 uses a set of ground stations which spans the globe, the periods during which measurements are not available to the filter are much shorter. In fact most of these outages are due to the minimum altitude constraint mentioned in Section 5.3. That is, there is a low-altitude portion of the highly-elliptical GTO during which no measurements are assumed to be available for orbit determination. Case E1 performs very well. The peak position error after the initial filter transient is significantly less than 0.5 km for this example, and the RMS error over days 8–14 is only about 35 m.

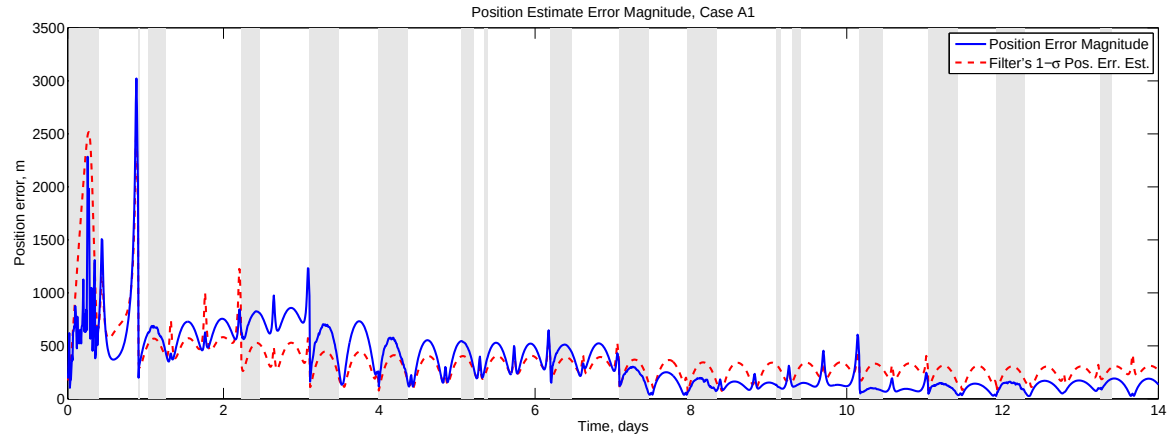
The orbit determination performance for the intermediate orbit, with periapsis halfway between GTO periapsis and GEO radius, is in general somewhat worse than for the GTO orbit. An example of case A2 is shown in Figure 5.5. This orbit is considerably higher than the GTO, and thus it is also slower. Data gaps are lengthened to nearly two days in this scenario. However, counter balancing these longer gaps are longer periods during which the satellite is visible to two or more ground stations.

When additional ground stations are added, to produce Case E2 shown in Figure 5.6, the filter's performance improves. There are very few data gaps, and they are short. The peak position estimation errors climb to only about two hundred meters. The RMS position estimation error for days 8–14 of this example is about 68 m, about twice that for Case E1.

The filter presented in this dissertation includes terms that allow unmodeled conservative and non-conservative forces to be absorbed into empirical acceleration parameters estimated on the fly. These terms allow for some amount of uncertainty to exist in the force model without causing problems with the filter. While the filter presented herein has not been applied to on-orbit data, nor was it meant to be in its present simplified form, the results of the simulated orbit determination scenarios should adequately indicate the levels of orbit de-

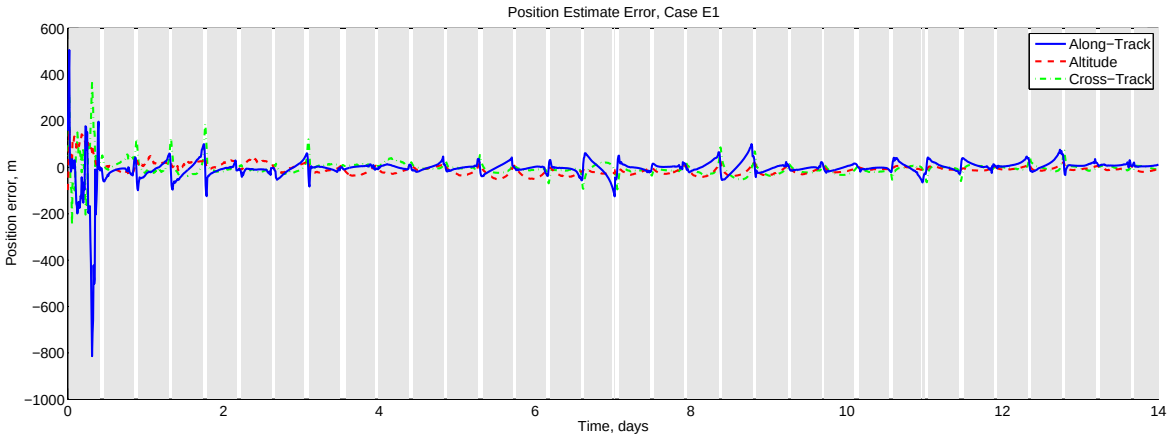


(a) Along-track (solid blue), altitude (dashed red), and cross-track (dashed-dotted green) position estimation errors are plotted versus time. The vertical grey bars mark the time spans during which the satellite is being observed by two or more ground stations.

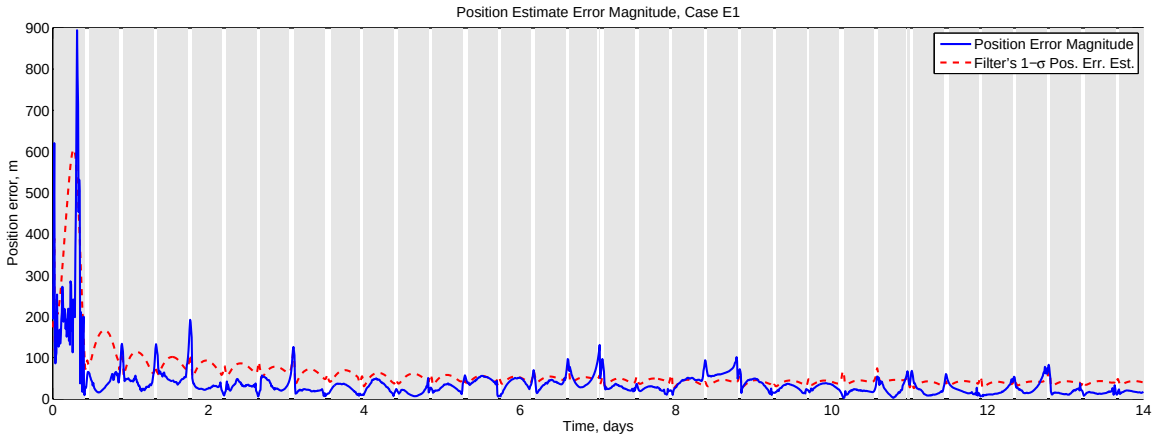


(b) The magnitude of the position estimation error (solid blue) and the filter's 1-σ estimate of the position estimation error (dashed red) are plotted versus time. The grey bars again mark times during which the filter is receiving measurements.

Figure 5.3 – Case A1 position estimation error components and magnitude. Case A1 uses the GTO orbit, six ground stations scattered through the contiguous United States, and 0.5 Hz carrier Doppler shift measurement error standard deviation.

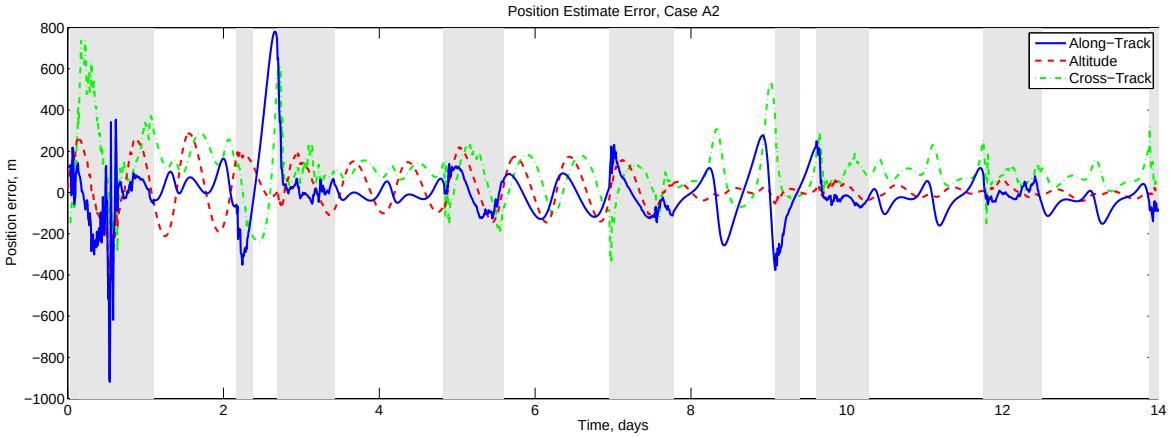


(a) Along-track (solid blue), altitude (dashed red), and cross-track (dashed-dotted green) position estimation errors are plotted versus time. The vertical grey bars mark the time spans during which the satellite is being observed by two or more ground stations.

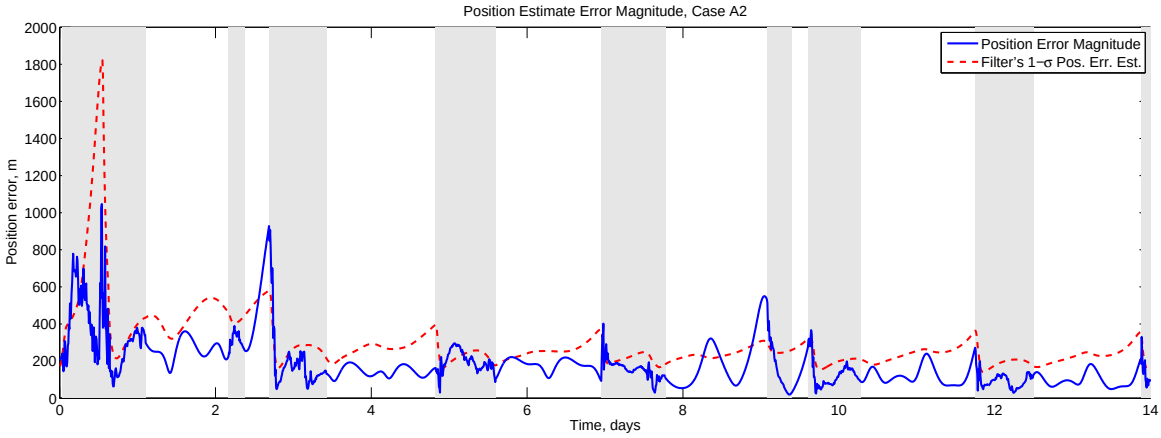


(b) The magnitude of the position estimation error (solid blue) and the filter's 1-σ estimate of the position estimation error (dashed red) are plotted versus time. The grey bars again mark times during which the filter is receiving measurements.

Figure 5.4 – Case E1 position estimation error components and magnitude. Case E1 uses the GTO orbit, all 17 ground stations scattered around the Earth, and 0.5 Hz carrier Doppler shift measurement error standard deviation.

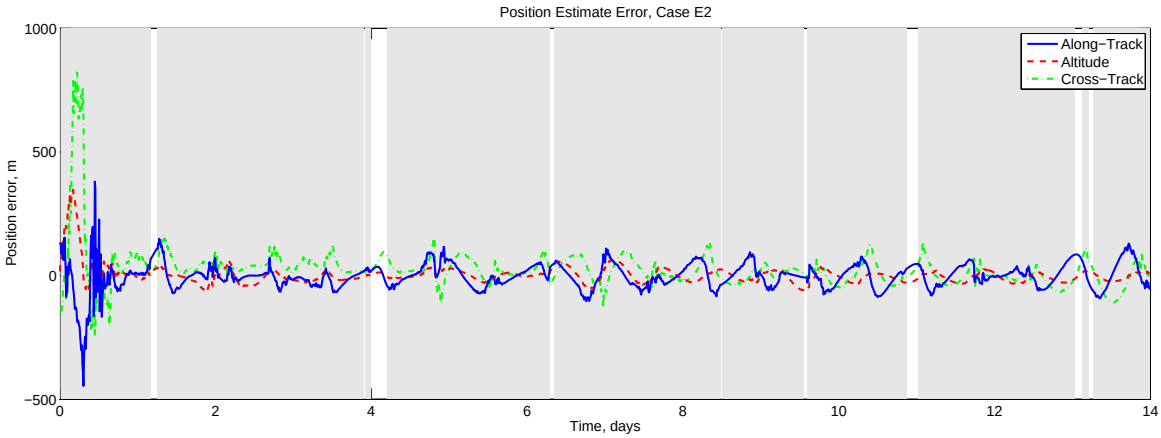


(a) Along-track (solid blue), altitude (dashed red), and cross-track (dashed-dotted green) position estimation errors are plotted versus time. The vertical grey bars mark the time spans during which the satellite is being observed by two or more ground stations.

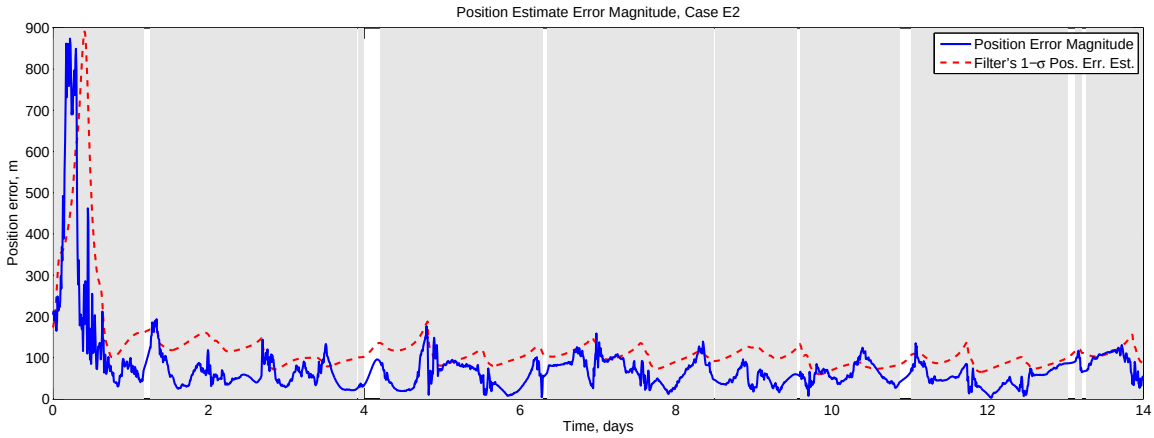


(b) The magnitude of the position estimation error (solid blue) and the filter's 1-σ estimate of the position estimation error (dashed red) are plotted versus time. The grey bars again mark times during which the filter is receiving measurements.

Figure 5.5 – Case A2 position estimation error components and magnitude. Case A2 uses the intermediate orbit, six ground stations scattered through the contiguous United States, and 0.5 Hz carrier Doppler shift measurement error standard deviation.



(a) Along-track (solid blue), altitude (dashed red), and cross-track (dashed-dotted green) position estimation errors are plotted versus time. The vertical grey bars mark the time spans during which the satellite is being observed by two or more ground stations.



(b) The magnitude of the position estimation error (solid blue) and the filter's 1-σ estimate of the position estimation error (dashed red) are plotted versus time. The grey bars again mark times during which the filter is receiving measurements.

Figure 5.6 – Case E2 position estimation error components and magnitude. Case E2 uses the intermediate orbit, all 17 ground stations scattered around the Earth, and 0.5 Hz carrier Doppler shift measurement error standard deviation.

termination accuracy possible given the measurements and geometry of the problem. When the orbit determination methods set out herein are applied to a real orbit determination problem, the satellite force model and measurement model must be enhanced to yield higher fidelity than they do currently. Even so, it is likely to be desirable to retain the constant and once-per-revolution empirical acceleration disturbance parameters in order to absorb any residual model errors.

CHAPTER 6

SUMMARY

6.1 Summary and Review

The goal of this dissertation is to provide a proof of concept and the implementation details of a new, proposed low-cost orbit determination method. The proposed method would enable a new class of commercial satellite mission by significantly reducing the cost of orbit determination during the orbit raising phase of the satellite's life. The proposed method uses passive measurements of the satellite's low-power, always-on normal downlink telemetry signal and uses them in an extended square root information filter to compute estimates of satellite position and velocity. The proposed measurements are measurements of telemetry signal carrier Doppler shift. These measurements contain information related to the relative motion between the satellite and the ground station. The measurements are contaminated with spacecraft clock frequency error and receiver clock frequency error. It is assumed that the receiver clock frequency error can be calibrated out using GPS measurements or through some other means. The satellite clock errors are removed by observing the satellite signal simultaneously from more than one location. The measurements are made through a matched filtering process that starts with a coarse grid search over signal parameter space and finishes with a non-linear optimization of the parameters found during the grid search. The measurements from a network of ground stations are then passed to a central orbit determination algorithm. The proposed method uses an extended square root information filter.

This dissertation started with the development of a signal model of a downlink telemetry signal that is being proposed for use in passive orbit determination. This type of signal is referred to in the satellite industry as the "normal downlink telemetry" as opposed to "dwell

telemetry" or another special telemetry signal. The model was specific to the signal investigated during the proof-of-concept orbit determination signal processing experiment which was subsequently described in Chapter 3. The development of the signal model proceeded by first setting forth a mathematical description of the signal, then developing approximations that simplified the parameter space of the signal.

The signal model was then applied in a coarse acquisition algorithm. The coarse acquisition algorithm was tested in an experiment in Gilbert, Arizona. RF samples of the downlink telemetry from an on-orbit satellite were simultaneously captured using a high-gain 13-meter antenna and a low-gain antenna. The coarse acquisition algorithm was applied to the signal samples from both antennas. The high-gain antenna provided some measure of "truth", while the low-gain antenna provided measurements with the fidelity expected during operation of the low-cost orbit determination method. The results of the proof-of-concept orbit determination signal processing experiment proved that it is possible to detect the signal at the expected levels of C/N_0. This was an open question prior to the experiment. It was unknown whether the signal could be detected at all at very low C/N_0.

A fine acquisition algorithm refined the results of the coarse acquisition and provided carrier Doppler measurements at reasonable levels of measurement noise. The variance of the measurement noise was determined theoretically using the Cramér-Rao Lower Bound. The difference in carrier Doppler shift between the high-SNR and low-SNR detections was taken to be a proxy for the measurement error. The magnitude of the measurement errors was found to be consistent with the theoretically determined measurement noise variance.

The next chapter of the dissertation demonstrated that a modestly-sized network of low-cost ground stations placed at strategic locations around the globe would be able to provide carrier Doppler measurements to a filter to enable real-time orbit determination. In fact, though the only orbit determination method investigated was real-time square root information filtering, the measurements used could be provided to any type of orbit determination

method capable of processing carrier Doppler measurements. The true achievable levels of position estimate error are currently unknown, but the simulation results suggest that it is possible to compute orbits with less than 100 m RMS position error using the proposed system. Achieving this level of accuracy will be a matter of thoughtful hardware and software development and careful modeling.

The overall cost to develop and deploy this orbit determination system may ultimately be less than the cost of traditional orbit determination for an extended orbit-raising campaign. The recurring cost of operating a network of the proposed ground stations is minimal—occasional inspections, power, and internet connection. This system has the potential to save millions of dollars per year or to enable new capability at low cost.

6.2 Next Steps

In order to pursue this system, a few steps should be taken. As mentioned earlier, the spacecraft engineering community is risk averse and extremely conservative. Accordingly, the development of this system should proceed in small, low-risk phases.

1. Develop prototype ground stations
2. Develop preliminary orbit and measurement models
3. Build two or more prototype ground stations
4. Perform orbit determination for an on-orbit satellite (or during orbit raising)

The development and build of prototype ground stations can be accomplished in very little time at very low cost. The hardware configuration for the proof-of-concept experiment described in Section 3.5.1 would serve as an adequate starting point. Some method

of providing highly accurate time stamps would be necessary, for instance, simultaneously sampling the GPS signal and time stamping the RF samples after processing the GPS data.

Before actual orbit determination can be accomplished, some work must be done on the orbit and measurement models. Those presented in this dissertation are not sufficiently high-fidelity. For instance, a full Earth gravity model should be used in the filter, solar and lunar gravitation should be accounted for, a drag model may be needed for low orbits or orbits with low perigee, and even solar radiation pressure will be important. One effect that has not been addressed is orbital maneuvers. These must be dealt with in a sensible way. For instance, an *a priori* thrust magnitude and direction could be applied at the correct time step, and a temporary state added to the filter to estimate a correction to the *a priori* level. The measurement model must be developed to take into account light time propagation delay of the signal, additional propagation delay due to passage through the neutral atmosphere, and if the frequency is low enough, the carrier phase advance caused by the ionosphere. It will be important to rigorously account for the differing measurement effectiveness times [34] between ground stations.

After building two prototype ground stations, initial orbit determination can be demonstrated. If the two ground stations are deployed in widely separated locations both in-view of an on-orbit satellite, a very rough orbit solution could be calculated using data from both of them. This could then be compared to the satellite operator's orbit solution.

If the orbits matched to within the bounds of the uncertainty, then further development of this system could easily be justified. From this point in the development, it would be a short step to deploying several ground stations and computing high-quality orbits.

Some additional considerations that may reduce project risk or increase the orbit determination accuracy are worth mentioning. It would simplify orbit determination if the subcarrier and the carrier were not independent. For the class of communication satellite investigated,

the carrier-generating clock runs independently of the subcarrier-generating clock. This is not problematic when the signal is observed using a high-gain parabolic antenna, because the carrier and subcarrier can be acquired independently in series. However, when observing using a low-gain antenna, the carrier and subcarrier may need to be acquired simultaneously. If the two clocks were synchronized, this process would be immensely simplified. Of the five non-linear signal parameters, (α_{Ak}, ω_{DAk}, α_{B0}, ω_{DB0}, and τ_{B0}) the two pertaining to the subcarrier frequency could be eliminated. This would reduce the search space from three dimensions to two dimensions.

Using higher-gain antennas provides higher-quality measurements but increases the complexity of the system. Depending on the antennas used, it may either increase or decrease the cost of a ground station. Using low-cost consumer antennas, such as 75-cm K_u band satellite television receiving antennas, may actually decrease the cost of a system, because the antennas are so inexpensive[1]. However, a higher-gain antenna must be pointed more accurately than a lower-gain antenna. In the extreme cases, a high-gain antenna must be dynamically steered to point directly at the satellite during the duration of the observation, whereas an omnidirectional antenna may simply be pointed at the sky. The additional cost and complexity of antenna pointing and steering may be prohibitive. However, there are less extreme possibilities. A medium gain antenna, like a consumer K_u band satellite television antenna, has a main beam that is broad enough that the antenna could be pointed to the location of the satellite at the center of an observation and left to dwell there over the duration of the observation. Alternatively, a phased array of lower-gain antennas could be used to steer the main beam of the array without any moving parts. In fact, a phased array could be used to observe multiple satellites simultaneously by steering the beam in post-processing. Another possibility would be to have one or more "master stations" that have higher-gain antennas. These master stations could provide the *a priori* information necessary to aid acquisition using low-gain antennas.

[1] A 30-inich K_u band satellite dish costs about \$40 at the time of writing.

REFERENCES

[1] Michel Baylocq. Private communication, 2014.

[2] Chris Hoeber. Private communication, 2014.

[3] Pratap Misra and Per Enge. *Global Positioning System: Signals, Measurements, and Performance*. Ganga-Jamuna Press, Lincoln, Massachusetts, 2nd edition, 2006. pp 410, 504, 505.

[4] Rick Voreck. Private communication, Dec 2012.

[5] International Telephone and Telegraph Corporation. *Reference Data For Radio Engineers*. Howard W Sams and Company, New York, 5th edition, 1972.

[6] A.J. Van Dierendonck. GPS Receivers. In B.W. Parkinson and J.J. Spilker, editors, *Global Positioning System: Theory and Applications*, number 163 in Progress in Astronautics and Aeronautics, pages 402–405. American Institute of Aeronautics and Astronautics, 1996.

[7] R. Tandra and A. Sahai. SNR Walls for Signal Detection. *IEEE Journal of Selected Topics in Signal Processing*, 2(1):4–17, Feb 2008.

[8] Mark L. Psiaki. FFT-based acquisition of GPS L2 civilian CM and CL signals. *Proceedings of the 17th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2004)*, pages 457–473, Long Beach, CA, September 2004.

[9] H. Vincent Poor. *An introduction to signal detection and estimation*. Springer-Verlag, New York, 1994. pp 5–39.

[10] Mark L. Psiaki. Block acquisition of weak GPS signals in a software receiver. *Proceedings of the 14th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2001)*, pages 2838–2850, Salt Lake City, UT, September 2001.

[11] Richard F. Voss. 1/f (flicker) noise: A brief review. *Proceedings of the 33rd Annual Symposium on Frequency Control*, pages 40–46, Atlantic City, New Jersey, May 1979.

[12] Ryan C. Dougherty and Mark L. Psiaki. Orbit determination using TDMA radio navigation data with implicit measurement times. *Proc. AIAA Guidance, Navigation, and Control Conf.*, Portland, OR, August 2011.

[13] Mark L. Psiaki. Ultra-precise orbit, clock, ionosphere, and troposphere estimation for iGPS. Unpublished technical report, June 2013.

[14] Code of Federal Regulations, Title 47, Part 25 – Satellite Communications, 2010.

[15] Recommendation ITU-R S.1716, Performance and availability objectives for fixed-satellite service telemetry, tracking and command systems, 2005.

[16] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer, New York, 1999. pp 250–262.

[17] Yaakov Bar-Shalom, X. Rong Li, and Thiagalingam Kirubarajan. *Estimation with Applications to Tracking and Navigation*. John Wiley and Sons, New York, New York, 2001. pp 109–115.

[18] J. C. Ries, C. K. Shum, and B. D. Tapley. Surface force modeling for precision orbit determination. *Environmental Effects on Spacecraft Positioning and Trajectories*, pages 111–124. American Geophysical Union, 2013.

[19] B. D. Tapley, J. C. Ries, G. W. Davis, R. J. Eanes, B. E. Schutz, C. K. Shum, M. M. Watkins, J. A. Marshall, R. S. Nerem, B. H. Putney, S. M. Klosko, S. B. Luthcke, D. Pavlis, R. G. Williamson, and N. P. Zelensky. Precision orbit determination for topex/poseidon. *Journal of Geophysical Research: Oceans*, 99(C12):24383–24404, 1994.

[20] W. I. Bertiger, Y. E. Bar-Sever, E. J. Christensen, E. S. Davis, J. R. Guinn, B. J. Haines, R. W. Ibanez-Meier, J. R. Jee, S. M. Lichten, W. G. Melbourne, R. J. Muellerschoen, T. N. Munson, Y. Vigue, S. C. Wu, T. P. Yunck, B. E. Schutz, P. A. M. Abusali, H. J. Rim, M. M. Watkins, and P. Willis. GPS precise tracking of TOPEX/POSEIDON: Results and implications. *Journal of Geophysical Research: Oceans*, 99(C12):24449–24464, 1994.

[21] R.A. Broucke and P.J. Cefola. On the equinoctial orbit elements. *Celestial Mechanics*, 5(3):303–310, 1972.

[22] Oliver Montenbruck and Eberhard Gill. *Satellite Orbits: Models, Methods and Applications*. Springer, Berlin, Heidelberg, and New York, 2005. p 55.

[23] Thomas P. Yunck. Orbit Determination. In B.W. Parkinson and J.J. Spilker, editors, *Global Positioning System: Theory and Applications*, number 164 in Progress in Astronautics and Aeronautics, page 561. American Institute of Aeronautics and Astronautics, 1996.

[24] Robert Grover Brown and Patrick Y. C. Hwang. *Introduction to Random Signals and*

Applied Kalman Filtering with MATLAB exercises and solutions. Wiley, New York, 3rd edition, November 1996. pp 428–432.

[25] Paul G. Kaminski, Jr. Arthur E. Bryson, and Stanley F. Schmidt. Discrete square root filtering: A survey of current techniques. *IEEE Transactions on Automatic Control*, AC-16(6):727–726, December 1971.

[26] Gerald J. Bierman. *Factorization Methods for Discrete Sequential Estimation*. Academic Press, New York, New York, 1977. pp 69–76, 115–122.

[27] A. L. Maldonado, M. Baylocq, and G. Hannan. Autonomous spacecraft navigation - Extended Kalman filter estimation of classical orbital parameters. In R. D. Culp and S. S. Parker, editors, *Guidance and Control Conference*, pages 9–18, 1984.

[28] Gerald J. Bierman. Square-root information filtering and smoothing for precision orbit determination. In Danny C. Sorensen and Roger J.-B. Wets, editors, *Algorithms and Theory in Filtering and Control*, volume 18 of *Mathematical Programming Studies*, pages 61–75. Springer Berlin Heidelberg, 1982.

[29] Mark L. Psiaki. Satellite orbit determination using a single-channel global positioning system receiver. *Journal of Guidance, Control, and Dynamics*, 25(1):137–144, 2002.

[30] S.R. Best. Distance-measurement error associated with antenna phase-center displacement in time-reference radio positioning systems. *Antennas and Propagation Magazine, IEEE*, 46(2):13–22, April 2004.

[31] Ralf Schmid, Peter Steigenberger, Gerd Gendt, Maorong Ge, and Markus Rothacher. Generation of a consistent absolute phase-center correction model for GPS receiver and satellite antennas. *Journal of Geodesy*, 81(12):781–798, 2007.

[32] Oliver Montenbruck, Miquel Garcia-Fernandez, Yoke Yoon, Steffen Schön, and Adrian Jäggi. Antenna phase center calibration for precise positioning of LEO satellites. *GPS Solutions*, 13(1):23–34, 2009.

[33] Adrian Jäggi, R. Dach, O. Montenbruck, U. Hugentobler, H. Bock, and G. Beutler. Phase center modeling for LEO GPS receiver antennas and its impact on precise orbit determination. *Journal of Geodesy*, 83(12):1145–1162, 2009.

[34] Shan Mohiuddin and Mark L. Psiaki. Continuous-time Kalman filtering with implicit discrete measurement times. *Proc. AIAA Guidance, Navigation, and Control Conference*, pages 4574–4607, Washington, D.C., August 2009.